

420 Rec'd PCT/PTO 08 OCT 1999

1

TITLE OF THE INVENTION**PCA3, PCA3 GENES, AND METHODS OF USE****FIELD OF THE INVENTION**

5 The present invention relates, in general, to a prostate cancer antigen, PCA3. In particular, the present invention relates to nucleic acid molecules coding for the PCA3 protein; purified PCA3 proteins and polypeptides; recombinant nucleic acid molecules; cells containing the recombinant nucleic acid molecules; antibodies having binding affinity
10 specifically to PCA3 proteins and polypeptides; hybridomas containing the antibodies; nucleic acid probes for the detection of nucleic acids encoding PCA3 proteins; a method of detecting nucleic acids encoding PCA3 proteins or polypeptides in a sample; kits containing nucleic acid probes or antibodies; bioassays using the nucleic acid sequence, protein or antibodies of this
15 invention to diagnose, assess, or prognose a mammal afflicted with prostate cancer; therapeutic uses; and methods of preventing prostate cancer in an animal.

BACKGROUND OF THE INVENTION

20 Prostate cancer is the most commonly diagnosed malignancy and the second leading cause of cancer-related deaths in the western male population. When this carcinoma has locally or distantly spread, no curative therapy can be offered. Therefore, efforts to control the disease (i.e., to decrease prostate cancer mortality) have focused on
25 increasing detection of the cancer while it is still locally confined and potentially curable. Studies aimed at the early detection of prostate cancer have demonstrated an appreciable increase in the detection of organ-confined potentially curable prostate cancers. However, it has not yet been demonstrated that the increased detection rate will decrease the prostate

cancer-specific mortality rates. On the other hand, there is also no evidence that early diagnosis will decrease the mortality rates. Both in the United States and in Europe, discussions on the efficacy and acceptability of screening programs, the issue of overdiagnosis and overtreatment and the chances that early treatment will lead to reduced prostate cancer morbidity and mortality, are still ongoing and make early detection of prostate cancer a controversial issue (Schröder, *Urology* 46: 6270 (1995)).

Measurements of serum concentrations of prostatic marker enzymes have recognized value in the clinical detection, diagnosis and management of prostate cancer. The two most widely used prostatic marker enzymes are prostatic acid phosphatase (PAP) and prostate-specific antigen (PSA). Normally, both enzymes are secreted from the prostatic epithelial cells into the seminal fluid, but in patients with prostatic disease they leak into the circulation, where they can be detected by means of immunological assays (Armbruster, *Clin. Che.* 39: 181-95 (1993)).

Prostatic acid phosphatase, one of the earliest serum markers for prostate, has an as yet undetermined function and is one of the most predominant protein components in human prostatic secretions. The use of PAP as a marker for prostatic tumors is complicated by the reported structural similarities between the prostate-specific acid phosphatase and the lysosomal acid phosphatase occurring in all tissues. Furthermore, there is a tendency towards lower PAP mRNA and protein levels in prostate cancer in comparison with benign prostatic hyperplasia (BPH). In recent years, PAP measurements were superseded by serum PSA measurements in the clinical management of prostate cancer.

Prostate-specific antigen was identified by several groups in the 1970's as a prostate-specific protein from the seminal fluid. In 1979, it was purified as an antigen from prostate cancer tissue. Further research showed that PSA is produced exclusively by the columnar epithelial cells of

the prostate and periureteral glands. Normal prostate epithelium and benign hyperplastic tissue actually produce more PSA mRNA and protein than does prostate cancer tissue. Furthermore, it was shown that loss of differentiation of prostatic carcinomas is associated with a decrease in the level of intraprostatic PSA.

Abnormalities in prostate architecture occurring as a result of prostatic disease lead to an increased leakage of PSA (and PAP) into the serum and make serum PSA measurements a marker for prostate cancer. Despite the fact that early studies have indicated that diagnostic PSA testing would be hampered by the fact that it lacked specificity in differentiating between BPH and prostate cancer, PSA testing was introduced in 1986 and revolutionized the management of patients with prostate cancer. Increased knowledge on the organ specificity of PSA and the relationship of elevated serum PSA levels to prostate disease as well as improvement of biopsy techniques and histological evaluation, led to a appreciation of the clinical value of PSA testing, a utility not yet achieved by any other (prostate) tumor marker. Cloning of the gene that encodes PSA revealed that it is a member of the human kallikrein gene family and resulted in the development of new approaches to the use of PSA as a marker: the very sensitive reverse transcriptase polymerase chain reaction (RT-PCR) method is used to detect extremely small numbers of malignant prostate cells in blood samples from prostate cancer patients and might provide a sensitive tool to identify patients with micrometastatic disease (Moreno *et al.*, *Cancer Res.* 52: 6110-12 (1992); and Katz *et al.*, *Urology* 43: 765-75 (1994)).

Prostate-specific membrane antigen (PSM) was originally identified using an antibody developed by immunizing mice with the membrane fraction of LNCaP human prostatic adenocarcinoma cells. Like PAP and PSA, PSM can be detected in normal prostate, BPH and prostate cancer and is absent from most other tissues. Also for PSM, RT-PCR studies

have been developed to detect circulating prostate cancer cells, however, further investigations are required to establish the usefulness of PSM as marker for prostatic cancer.

5 In summary, PSA is currently recognized as the premier
marker for prostatic cancer, being useful for screening selected populations
of patients with symptoms indicative of prostate cancer and for monitoring
patients after therapy, especially after surgical prostatectomy (measurable
levels of PSA indicate residual disease or metastasis and increasing PSA
concentrations indicate recurrent disease). The significant weaknesses of
10 PSA as a tumor marker are that (1) PSA is not able to always distinguish
prostate cancer from BPH; and (2) that its expression decreases with loss of
differentiation of carcinomas.

In view of the fact that advanced prostate cancer remains
a life threatening disease reaching a very significant proportion of the male
15 population, there remains a need for the development of new treatment and
diagnostic modalities for (late stage) prostate cancer.

The present invention seeks to meet these and other
needs.

20 The present description refers to a number of documents,
the content of which is herein incorporated by reference.

SUMMARY OF THE INVENTION

The invention provides, in general, isolated nucleic acid
molecules coding for PCA3 or fragments thereof.

25 The invention further provides purified polypeptides
encoding PCA3 or an epitope binding portion thereof.

The invention also provides nucleic acids for the specific
detection of the presence of nucleic acids encoding PCA3 proteins or
polypeptides in a sample.

The invention further provides a method of detecting nucleic acid encoding PCA3 in a sample.

The invention also provides a kit for detecting the presence of nucleic acid encoding PCA3 in a sample.

5 The invention further provides a recombinant nucleic acid molecule comprising, 5' to 3', a promoter effective to initiate transcription in a host cell and the above-described isolated nucleic acid molecule.

The invention also provides a recombinant nucleic acid molecule comprising a vector and the above-described isolated nucleic acid molecule.

10 The invention further provides an antisense PCA3 nucleic acid molecule.

The invention also provides a cell that contains the above-described recombinant nucleic acid molecule.

15 The invention further provides a non-human organism that contains the above-described recombinant nucleic acid molecule.

The invention also provides an antibody having binding affinity specifically to PCA3 or an epitope-bearing portion thereof.

20 The invention further provides a method of detecting PCA3 in a sample.

The invention also provides a method of measuring the amount of PCA3 in a sample.

25 The invention in addition provides immunogenic reagents to induce protection against PCA3-expressing prostate cancer cells. Preferably, such immunogenic reagents are polypeptides encoding PCA3, an antigenic portion thereof, fusion proteins encoding PCA3 or fusion protein encoding antigenic portions of PCA3. In such an embodiment, these immunogenic reagents would function as vaccine agents.

The invention further provides a method of detecting antibodies having binding affinity specifically to PCA3.

The invention further provides a diagnostic kit comprising a first container means containing the above-described antibody, and a
5 second container means containing a conjugate comprising a binding partner of the monoclonal antibody and a label.

The invention also provides a hybridoma which produces the above-described monoclonal antibody.

The invention further provides diagnostic methods for
10 human disease, in particular, prostate cancer. Preferably, a method of diagnosing the presence or predisposition to develop prostate cancer in a patient is provided herein.

The invention also provides methods for therapeutic uses involving all or part of (1) a nucleic acid sequence encoding PCA3, (2)
15 antisense PCA3 nucleic acid molecules, (2) PCA3 protein, or (4) PCA3 antibodies.

Further objects and advantages of the present invention will be clear from the description that follows.

20 DEFINITIONS

In the description that follows, a number of terms used in recombinant DNA (rDNA) technology are extensively utilized. In order to provide a clear and consistent understanding of the specification and claims, including the scope to be given such terms, the following definitions are
25 provided.

Isolated Nucleic Acid Molecule. An "isolated nucleic acid molecule", as is generally understood and used herein, refers to a polymer of nucleotides, and includes but should not be limited to DNA and RNA. The "isolated" nucleic acid molecule is purified from its natural *in vivo* state.

Recombinant DNA. Any DNA molecule formed by joining DNA segments from different sources and produced using recombinant DNA technology (aka. molecular genetic engineering).

5 DNA Segment. A DNA segment, as is generally understood and used herein, refers to a molecule comprising a linear stretch of nucleotides wherein the nucleotides are present in a sequence that can encode, through the genetic code, a molecule comprising a linear sequence of amino acid residues that is referred to as a protein, a protein fragment or a polypeptide.

10 Gene. A DNA sequence related to a single polypeptide chain or protein, and as used herein includes the 5' and 3' untranslated ends. The polypeptide can be encoded by a full-length sequence or any portion of the coding sequence, so long as the functional activity of the protein is retained.

15 Complementary DNA (cDNA). Recombinant nucleic acid molecules synthesized by reverse transcription of messenger RNA ("mRNA").

20 Structural Gene. A DNA sequence that is transcribed into mRNA that is then translated into a sequence of amino acids characteristic of a specific polypeptide.

25 Restriction Endonuclease. A restriction endonuclease (also restriction enzyme) is an enzyme that has the capacity to recognize a specific base sequence (usually 4, 5, or 6 base pairs in length) in a DNA molecule, and to cleave the DNA molecule at every place where this sequence appears. For example, *EcoRI* recognizes the base sequence GAATTC/CTTAAG.

Restriction Fragment. The DNA molecules produced by digestion with a restriction endonuclease are referred to as restriction

fragments. Any given genome can be digested by a particular restriction endonuclease into a discrete set of restriction fragments.

Agarose Gel Electrophoresis. To detect a polymorphism in the length of restriction fragments, an analytical method for fractionating double-stranded DNA molecules on the basis of size is required. The most commonly used technique (though not the only one) for achieving such a fractionation is agarose gel electrophoresis. The principle of this method is that DNA molecules migrate through the gel as though it were a sieve that retards the movement of the largest molecules to the greatest extent and the movement of the smallest molecules to the least extent. Note that the smaller the DNA fragment, the greater the mobility under electrophoresis in the agarose gel.

The DNA fragments fractionated by agarose gel electrophoresis can be visualized directly by a staining procedure if the number of fragments included in the pattern is small. The DNA fragments of genomes can be visualized successfully. However, most genomes, including the human genome, contain far too many DNA sequences to produce a simple pattern of restriction fragments. For example, the human genome is digested into approximately 1,000,000 different DNA fragments by *EcoRI*. In order to visualize a small subset of these fragments, a methodology referred to as the Southern hybridization procedure can be applied.

Southern Transfer Procedure. The purpose of the Southern transfer procedure (also referred to as blotting) is to physically transfer DNA fractionated by agarose gel electrophoresis onto a nitrocellulose filter paper or another appropriate surface or method, while retaining the relative positions of DNA fragments resulting from the fractionation procedure. The methodology used to accomplish the transfer from agarose gel to nitrocellulose involves drawing the DNA from the gel into the nitrocellulose paper by capillary action.

5 Nucleic Acid Hybridization. Nucleic acid hybridization depends on the principle that two single-stranded nucleic acid molecules that have complementary base sequences will reform the thermodynamically favored double-stranded structure if they are mixed under the proper conditions. The double-stranded structure will be formed between two complementary single-stranded nucleic acids even if one is immobilized on a nitrocellulose filter. In the Southern hybridization procedure, the latter situation occurs. As noted previously, the DNA of the individual to be tested is digested with a restriction endonuclease, fractionated by agarose gel electrophoresis, converted to the single-stranded form, and transferred to nitrocellulose paper, making it available for reannealing to the hybridization probe. Examples of hybridization conditions can be found in Ausubel, F.M. *et al.*, *Current protocols in Molecular Biology*, John Wiley & Sons, Inc., New York, NY (1989). A nitrocellulose filter is incubated overnight at 68°C with labeled probe in a solution containing 50% formamide, high salt (either 5x SSC[20X: 3M NaCl/0.3M trisodium citrate] or 5X SSPE [20X: 3.6M NaCl/0.2M NaH₂PO₄/0.02M EDTA, pH 7.7]), 5X Denhardt's solution, 1% SDS, and 100 µg/ml denatured salmon sperm DNA. This is followed by several washes in 0.2X SSC/0.1% SDS at a temperature selected based on the desired stringency: room temperature (low stringency), 42°C (moderate stringency) or 68°C (high stringency). The temperature selected is determined based on the melting temperature (T_m) of the DNA hybrid.

Hybridization Probe. To visualize a particular DNA sequence in the Southern hybridization procedure, a labeled DNA molecule or hybridization probe is reacted to the fractionated DNA bound to the nitrocellulose filter. The areas on the filter that carry DNA sequences complementary to the labeled DNA probe become labeled themselves as a consequence of the reannealing reaction. The areas of the filter that exhibit

such labeling are visualized. The hybridization probe is generally produced by molecular cloning of a specific DNA sequence.

Oligonucleotide or Oligomer. A molecule comprised of two or more deoxyribonucleotides or ribonucleotides, preferably more than three. Its exact size will depend on many factors, which in turn depend on the ultimate function or use of the oligonucleotide. An oligonucleotide can be derived synthetically or by cloning.

Sequence Amplification. A method for generating large amounts of a target sequence. In general, one or more amplification primers are annealed to a nucleic acid sequence. Using appropriate enzymes, sequences found adjacent to, or in between the primers are amplified.

Amplification Primer. An oligonucleotide which is capable of annealing adjacent to a target sequence and serving as an initiation point for DNA synthesis when placed under conditions in which synthesis of a primer extension product which is complementary to a nucleic acid strand is initiated.

Antisense nucleic acid molecule. An "antisense nucleic acid molecule" refers herein to a molecule capable of forming a stable duplex or triplex with a portion of its targeted nucleic acid sequence (DNA or RNA). The use of antisense nucleic acid molecules and the design and modification of such molecules is well known in the art as described for example in WO 96/32966, WO 96/11266, WO 94/15646, WO 93/08845, and USP 5,593,974. Antisense nucleic acid molecules according to the present invention can be derived from the nucleic acid sequences of the present invention and modified in accordance to well known methods. For example, some antisense molecules can be designed to be more resistant to degradation, to increase their affinity to their targeted sequence, to affect their transport to chosen cell types or cell compartments, and/or to enhance their lipid solubility

by using nucleotide analogs and/or substituting chosen chemical fragments thereof, as commonly known in the art.

Vector. A plasmid or phage DNA or other DNA sequence into which DNA can be inserted to be cloned. The vector can replicate autonomously in a host cell, and can be further characterized by one or a small number of endonuclease recognition sites at which such DNA sequences can be cut in a determinable fashion and into which DNA can be inserted. The vector can further contain a marker suitable for use in the identification of cells transformed with the vector. Markers, for example, are tetracycline resistance or ampicillin resistance. The words "cloning vehicle" are sometimes used for "vector."

Expression. Expression is the process by which a structural gene produces a polypeptide. It involves transcription of the gene into mRNA, and the translation of such mRNA into polypeptide(s).

Expression Vector. A vector or vehicle similar to a cloning vector but which is capable of expressing a gene which has been cloned into it, after transformation into a host. The cloned gene is usually placed under the control of (i.e., operably linked to) certain control sequences such as promoter sequences.

Expression control sequences will vary depending on whether the vector is designed to express the operably linked gene in a prokaryotic or eukaryotic host and can additionally contain transcriptional elements such as enhancer elements, termination sequences, tissue-specificity elements, and/or translational initiation and termination sites.

Functional Derivative. A "functional derivative" of a sequence, either protein or nucleic acid, is a molecule that possesses a biological activity (either functional or structural) that is substantially similar to a biological activity of the protein or nucleic acid sequence. A functional derivative of a protein can contain post-translational modifications such as

covalently linked carbohydrate, depending on the necessity of such modifications for the performance of a specific function. The term "functional derivative" is intended to include the "fragments," "segments," "variants," "analogs," or "chemical derivatives" of a molecule.

5 As used herein, a molecule is said to be a "chemical derivative" of another molecule when it contains additional chemical moieties not normally a part of the molecule. Such moieties can improve the molecule's solubility, absorption, biological half life, and the like. The moieties can alternatively decrease the toxicity of the molecule, eliminate or attenuate
10 any undesirable side effect of the molecule, and the like. Moieties capable of mediating such effects are disclosed in *Remington's Pharmaceutical Sciences* (1980). Procedures for coupling such moieties to a molecule are well known in the art.

Variant. A "variant" of a protein or nucleic acid is meant to
15 refer to a molecule substantially similar in structure and biological activity to either the protein or nucleic acid. Thus, provided that two molecules possess a common activity and can substitute for each other, they are considered variants as that term is used herein even if the composition or secondary, tertiary, or quaternary structure of one of the molecules is not identical to that
20 found in the other, or if the amino acid or nucleotide sequence is not identical.

Allele. An "allele" is an alternative form of a gene occupying a given locus on the chromosome.

Mutation. A "mutation" is any detectable change in the genetic material which can be transmitted to daughter cells and possibly even
25 to succeeding generations giving rise to mutant cells or mutant individuals. If the descendants of a mutant cell give rise only to somatic cells in multicellular organisms, a mutant spot or area of cells arises. Mutations in the germ line of sexually reproducing organisms can be transmitted by the gametes to the next generation resulting in an individual with the new mutant

condition in both its somatic and germ cells. A mutation can be any (or a combination of) detectable, unnatural change affecting the chemical or physical constitution, mutability, replication, phenotypic function, or recombination of one or more deoxyribonucleotides; nucleotides can be added, deleted, substituted for, inverted, or transposed to new positions with and without inversion. Mutations can occur spontaneously and can be induced experimentally by application of mutagens. A mutant variation of a nucleic acid molecule results from a mutation. A mutant polypeptide can result from a mutant nucleic acid molecule.

Species. A "species" is a group of actually or potentially interbreeding natural populations. A species variation within a nucleic acid molecule or protein is a change in the nucleic acid or amino acid sequence that occurs among species and can be determined by DNA sequencing of the molecule in question.

Polyacrylamide Gel Electrophoresis (PAGE). The most commonly used technique (though not the only one) for achieving a fractionation of polypeptides on the basis of size is polyacrylamide gel electrophoresis. The principle of this method is that polypeptide molecules migrate through the gel as though it were a sieve that retards the movement of the largest molecules to the greatest extent and the movement of the smallest molecules to the least extent. Note that the smaller the polypeptide fragment, the greater the mobility under electrophoresis in the polyacrylamide gel. Both before and during electrophoresis, the polypeptides typically are continuously exposed to the detergent sodium dodecyl sulfate (SDS), under which conditions the polypeptides are denatured. Native gels are run in the absence of SDS. The polypeptides fractionated by polyacrylamide gel electrophoresis can be visualized directly by a staining procedure if the number of polypeptide components is small.

Western Transfer Procedure. The purpose of the Western transfer procedure (also referred to as blotting) is to physically transfer polypeptides fractionated by polyacrylamide gel electrophoresis onto a nitrocellulose filter paper or another appropriate surface or method, while retaining the relative positions of polypeptides resulting from the fractionation procedure. The blot is then probed with an antibody that specifically binds to the polypeptide of interest.

Purified. A "purified" protein or nucleic acid is a protein or nucleic acid that has been separated from a cellular component. "Purified" proteins or nucleic acids have been purified to a level of purity not found in nature.

Substantially Pure. A "substantially pure" protein or nucleic acid is a protein or nucleic acid preparation that is lacking in all other cellular components.

BRIEF DESCRIPTION OF THE DRAWINGS

Having thus generally described the invention, reference will now be made to the accompanying drawings, showing by way of illustration a preferred embodiment thereof, and in which:

Figure 1 shows the genomic structure of the PCA3 gene.

Figure 2A-B, (A) a PCA3 cDNA structure; (B) a PCA3 nucleotide and amino acid cDNA sequence (SEQ ID NOs:1 and 2).

Figure 3A-B is a schematic representation comparing cDNA clones pMB9 and λ DD3.6.

Figure 4 shows the location of a transcription start site (TSS) of PCA3. The transcription start site was determined by primer extension (PE), S1-nuclease mapping (S1) and 5' Rapid Amplification of cDNA Ends (RACE) assays.

Figure 5A-B (A) a PCA3 cDNA structure; (B) a PCA3 nucleotide and amino acid cDNA sequence (SEQ ID NOs:6 and 7); putative poly-adenylation signals are underlined.

Other objects, advantages and features of the present invention will become more apparent upon reading of the following non-restrictive description of preferred embodiments with reference to the accompanying drawings which are exemplary and should not be interpreted as limiting the scope of the present invention.

10 DESCRIPTION OF THE PREFERRED EMBODIMENT

For purposes of clarity of disclosure, and not by way of limitation, the detailed description of the invention is divided into the following subsections:

- I. Isolated Nucleic Acid Molecules Coding for PCA3 Polypeptides.
- 15 II. Purified PCA3 Polypeptides.
- III. A Nucleic Acid Probe for the Specific Detection of PCA3 Nucleic Acid.
- IV. A Method of Detecting the Presence of PCA3 Nucleic Acid in a Sample.
- V. A Kit for Detecting the Presence of PCA3 Nucleic Acid in a Sample.
- VI. DNA Constructs Comprising a PCA3 Nucleic Acid Molecule and Cells
- 20 Containing These Constructs.
- VII. An Antibody Having Binding Affinity to a PCA3 Polypeptide and a Hybridoma Containing the Antibody.
- VIII. A Method of Detecting a PCA3 Polypeptide or Antibody in a Sample.
- IX. A Diagnostic Kit Comprising a PCA3 Protein or Antibody.
- 25 X. Diagnostic Screening
- XI. Therapeutic Treatments
- XII. Transgenic PCA3 Non-human Animals

I. Isolated Nucleic Acid Molecules Coding for PCA3 Polypeptides

In one embodiment, the present invention relates to isolated (purified) PCA3 nucleic acid molecules. Preferably, the PCA3 nucleic acid molecule comprises a polynucleotide sequence at least 90% identical (more preferably, 95%, 96%, 97%, 98%, 99% or 100% identical) to a sequence selected from the group consisting of:

- (a) a nucleotide sequence encoding a PCA3 polypeptide comprising the complete amino acid sequence in SEQ ID NO:2 or SEQ ID NO:7;
- (b) a nucleotide sequence encoding a PCA3 polypeptide comprising the complete amino acid sequence encoded by the polynucleotide clone contained in the deposit at the Centraal voor Schimmelcultures as accession number CBS 682.97;
- (c) a nucleotide sequence encoding a PCA3 polypeptide comprising the complete amino acid sequence encoded by the polynucleotide clone contained in the deposit at the Centraal voor Schimmelcultures as accession number CBS 100521; and
- (d) a nucleotide sequence complementary to any of the nucleotide sequences in (a), (b), or (c).

pMB9 is a PCA3 cDNA clone which contains exons 1, 2, 3, 4a and 4b of the PCA3 gene. pMB9 was deposited at the Phabagen Collection, University of Utrecht, Padualaan 8, 3584 CH Utrecht (which is a division of the Centraalbureau voor Schimmelcultures, Oosterstratt1, Postbus 273, 3740 AG Baarn) under the regulations of the Budapest Treaty on April 10, 1997 as accession number CBS 682.97.

ADD3.6 is a PCA3 cDNA clone which contains exons 3, 4a, 4b, 4c, and 4d. ADD3.6 was deposited at the Phabagen Collection, University of Utrecht, Padualaan 8, 3584 CH Utrecht (which is a division of the Centraalbureau voor Schimmelcultures, Oosterstratt1, Postbus 273, 3740 AG

Baarn) under the regulations of the Budapest Treaty on March 27, 1998 as accession number CBS 100521.

In one preferred embodiment, the isolated nucleic acid molecule comprises a PCA3 nucleotide sequence with greater than 90% identity or similarity to the nucleotide sequence present in SEQ ID NO:1 (preferably greater than 95%, 96%, 97%, 98%, 99% or 100%). In another preferred embodiment, the isolated nucleic acid molecule comprises the PCA3 coding sequence present in SEQ ID NO:1. In another embodiment, the isolated nucleic acid molecule encodes the PCA3 amino acid sequence present in SEQ ID NO:2 or SEQ ID NO:7. In yet another embodiment, the isolated nucleic acid molecule comprises a PCA3 nucleotide sequence with greater than 90% identity or similarity to the nucleotide sequence present in SEQ ID NO:6 (preferably greater than 95%, 96%, 97%, 98%, 99% or 100%). In another preferred embodiment, the isolated nucleic acid molecule comprises the PCA3 coding sequence present in SEQ ID NO:6.

Also included within the scope of this invention are isolated nucleic acids comprising cDNA splice variants of PCA3 or polynucleotide sequences which are at least 90% identical thereto, preferably at least 95% identical thereto. In view of the fact that virtually all combinations of exons are possible, non-limiting examples of such splice variants include isolated PCA3 nucleic acids comprising exons 1, 2, 3, 4a and 4b (SEQ ID NO:1); exons 1, 3, 4a, 4b, and 4c (SEQ ID NO:3 and region 4c which is contiguous to region 4b, see Figure 1); exons 1, 3, 4a, 4b, 4c, and 4d (SEQ ID NO:3 and region 4c which is contiguous to region 4b and region 4d which is contiguous to region 4c, see Figure 1); exons 1, 3, 4a, and 4b (SEQ ID NO:3); exons 1, 3, and 4a (SEQ ID NO:4); exons 1, 2, 3, 4a, 4b, 4c, and 4d (SEQ ID NO:6). Preferably, the PCA3 nucleic acid molecule comprises a polynucleotide sequence at least 90% identical (more preferably, 95%, 96%, 97%, 98%, 99% or 100% identical) to one of the above-described splice variants.

Also included within the scope of this invention are the functional equivalents of the herein-described isolated nucleic acid molecules and derivatives thereof. For example, the nucleic acid sequences depicted in SEQ ID NO:1 or SEQ ID NO:6 can be altered by substitutions, additions or deletions that provide for functionally equivalent molecules. Due to the degeneracy of nucleotide coding sequences, other DNA sequences which encode substantially the same amino acid sequence as depicted in SEQ ID NO:2 and SEQ ID NO:7 can be used in the practice of the present invention. These include but are not limited to nucleotide sequences comprising all or portions of PCA3 nucleic acid depicted in SEQ ID NO:1, 3, 4 or 6, which are altered by the substitution of different codons that encode a functionally equivalent amino acid residue within the sequence.

In addition, the nucleic acid sequence can comprise a nucleotide sequence which results from the addition, deletion or substitution of at least one nucleotide to the 5'-end and/or the 3'-end of the nucleic acid formula shown in SEQ ID NO:1, 3, 4, or 6 or a derivative thereof. Any nucleotide or polynucleotide can be used in this regard, provided that its addition, deletion or substitution does not substantially alter the amino acid sequence of SEQ ID NO:2, or SEQ ID NO:7 which is encoded by the nucleotide sequence. Moreover, the nucleic acid molecule of the present invention can, as necessary, have restriction endonuclease recognition sites added to its 5'-end and/or 3'-end. All variations of the nucleotide sequence of the PCA3 nucleotide coding sequence and fragments thereof permitted by the genetic code are, therefore, included in this invention.

Further, it is possible to delete codons or to substitute one or more codons by codons other than degenerate codons to produce a structurally modified polypeptide, but one which has substantially the same utility or activity of the polypeptide produced by the unmodified nucleic acid molecule. As recognized in the art, the two polypeptides are functionally

equivalent, as are the two nucleic acid molecules which give rise to their production, even though the differences between the nucleic acid molecules are not related to degeneracy of the genetic code.

5 **A. Isolation of Nucleic Acid**

In one aspect of the present invention, isolated nucleic acid molecules coding for polypeptides having amino acid sequences corresponding to PCA3 are provided. In particular, the nucleic acid molecule can be isolated from a biological sample containing PCA3 RNA or DNA.

10 The nucleic acid molecule can be isolated from a biological sample containing PCA3 RNA using the techniques of cDNA cloning and subtractive hybridization. The nucleic acid molecule can also be isolated from a cDNA library using a homologous probe.

15 The nucleic acid molecule can be isolated from a biological sample containing genomic DNA or from a genomic library. Suitable biological samples include, but are not limited to, whole organisms, organs, tissues, blood and cells. The method of obtaining the biological sample will vary depending upon the nature of the sample.

20 One skilled in the art will realize that genomes can be subject to slight allelic variations between individuals. Therefore, the isolated nucleic acid molecule is also intended to include allelic variations, so long as the sequence is a functional derivative of the PCA3 coding sequence. When a PCA3 allele does not encode the identical sequence to that found in SEQ ID NO:1 or 6, it can be isolated and identified as PCA3 using the same
25 techniques used herein, and especially PCR techniques to amplify the appropriate gene with primers based on the sequences disclosed herein.

One skilled in the art will realize that organisms other than humans will also contain PCA3 genes (for example, eukaryotes; more specifically, mammals, birds, fish, and plants; more specifically, gorillas,

rhesus monkeys, and chimpanzees). The invention is intended to include, but not be limited to, PCA3 nucleic acid molecules isolated from the above-described organisms.

5 **B. Synthesis of Nucleic Acid**

Isolated nucleic acid molecules of the present invention are also meant to include those chemically synthesized. For example, a nucleic acid molecule with the nucleotide sequence which codes for the expression product of a PCA3 gene can be designed and, if necessary, divided into
10 appropriate smaller fragments. Then an oligomer which corresponds to the nucleic acid molecule, or to each of the divided fragments, can be synthesized. Such synthetic oligonucleotides can be prepared, for example, by the triester method of Matteucci *et al.*, *J. Am. Chem. Soc.* 103:3185-3191 (1981) or by using an automated DNA synthesizer.

15 An oligonucleotide can be derived synthetically or by cloning. If necessary, the 5'-ends of the oligomers can be phosphorylated using T4 polynucleotide kinase. Kinasing of single strands prior to annealing or for labeling can be achieved using an excess of the enzyme. If kinasing is for the labeling of probe, the ATP can contain high specific activity
20 radioisotopes. Then, the DNA oligomer can be subjected to annealing and ligation with T4 ligase or the like.

II. Purified PCA3 Polypeptides

In another embodiment, the present invention relates to a
25 purified polypeptide (preferably, substantially pure) having an amino acid sequence corresponding to PCA3, or a functional derivative thereof. In a preferred embodiment, the polypeptide has the amino acid sequence set forth in SEQ ID NO:2 or 7 or mutant or species variation thereof, or at least 80% identity or at least 90% similarity thereof (preferably, at least 90%, 95%,

96%, 97%, 98%, or 99% identity or at least 95%, 96%, 97%, 98%, or 99% similarity thereof), or at least 6 contiguous amino acids thereof (preferably, at least 10, 15, 20, 25, or 50 contiguous amino acids thereof).

In a preferred embodiment, the invention relates to PCA3 epitopes. The epitope of these polypeptides is an immunogenic or antigenic epitope. An immunogenic epitope is that part of the protein which elicits an antibody response when the whole protein is the immunogen. An antigenic epitope is a fragment of the protein which can elicit an antibody response. Methods of selecting antigenic epitope fragments are well known in the art. See, Sutcliffe *et al.*, *Science* 219:660-666 (1983). Antigenic epitope-bearing peptides and polypeptides of the invention are useful to raise an immune response that specifically recognizes the polypeptides. Antigenic epitope-bearing peptides and polypeptides of the invention comprise at least 7 amino acids (preferably, 9, 10, 12, 15 or 20 amino acids) of the proteins of the invention. An example of a antigenic peptide is HTQEAQKEAQR (SEQ ID NO:5).

Amino acid sequence variants of PCA3 can be prepared by mutations in the DNA. Such variants include, for example, deletions from, or insertions or substitutions of, residues within the amino acid sequence shown in SEQ ID NO:2 or 7. Any combination of deletion, insertion, and substitution can also be made to arrive at the final construct, provided that the final construct possesses the desired activity.

While the site for introducing an amino acid sequence variation is predetermined, the mutation *per se* need not be predetermined. For example, to optimize the performance of a mutation at a given site, random mutagenesis can be conducted at the target codon or region and the expressed PCA3 variants screened for the optimal combination of desired activity. Techniques for making substitution mutations at predetermined sites

in DNA having a known sequence are well known, for example, site-specific mutagenesis.

Preparation of a PCA3 variant in accordance herewith is preferably achieved by site-specific mutagenesis of DNA that encodes an earlier prepared variant or a nonvariant version of the protein. Site-specific mutagenesis allows the production of PCA3 variants through the use of specific oligonucleotide sequences that encode the DNA sequence of the desired mutation. In general, the technique of site-specific mutagenesis is well known in the art, as exemplified by publications such as Adelman *et al.*, *DNA* 2:183 (1983) and Ausubel *et al.* "Current Protocols in Molecular Biology", J. Wiley & Sons, NY, NY, 1996.

Amino acid sequence deletions generally range from about 1 to 30 residues, more preferably 1 to 10 residues, and typically are contiguous.

Amino acid sequence insertions include amino and/or carboxyl-terminal fusions of from one residue to polypeptides of essentially unrestricted length, as well as intrasequence insertions of single or multiple amino acid residues. Intrasequence insertions (i.e., insertions within the complete PCA3 sequence) can range generally from about 1 to 10 residues, more preferably 1 to 5.

The third group of variants are those in which at least one amino acid residue in the PCA3 molecule, and preferably, only one, has been removed and a different residue inserted in its place. Such substitutions preferably are made in accordance with the following Table 1 when it is desired to modulate finely the characteristics of PCA3.

TABLE 1

	<u>Original Residue</u>	<u>Exemplary Substitutions</u>
	Ala	gly; ser
	Arg	lys
5	Asn	gln; his
	Asp	glu
	Cys	ser
	Gln	asn
	Glu	asp
10	Gly	ala; pro
	His	asn; gln
	Ile	leu; val
	Leu	ile; val
	Lys	arg; gln; glu
15	Met	leu; tyr; ile
	Phe	met; leu; tyr
	Ser	thr
	Thr	ser
	Trp	tyr
20	Tyr	trp; phe
	Val	ile; leu

Substantial changes in functional or immunological identity are made by selecting substitutions that are less conservative than those in Table 1, i.e., selecting residues that differ more significantly in their effect on maintaining (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a sheet or helical conformation, (b) the charge or hydrophobicity of the molecule at the target site, or (c) the bulk of the side chain. The substitutions that in general are expected are those in which (a) glycine and/or proline is substituted by another amino acid or is deleted or inserted; (b) a hydrophilic residue, e.g., seryl or threonyl, is substituted for (or by) a hydrophobic residue, e.g., leucyl, isoleucyl, phenylalanyl, valyl, or alanyl; (c) a cysteine residue is substituted for (or by) any other residue; (d) a residue having an electropositive side chain, e.g., lysyl, arginyl, or histidyl, is substituted for (or by) a residue having an electronegative charge, e.g.,

glutamyl or aspartyl; or (e) a residue having a bulky side chain, e.g., phenylalanine, is substituted for (or by) one not having such a side chain, e.g., glycine.

Some deletions and insertions, and substitutions are not expected to produce radical changes in the characteristics of PCA3. However, when it is difficult to predict the exact effect of the substitution, deletion, or insertion in advance of doing so, one skilled in the art will appreciate that the effect will be evaluated by routine screening assays. For example, a variant typically is made by site-specific mutagenesis of the native PCA3 encoding-nucleic acid, expression of the variant nucleic acid in recombinant cell culture, and, optionally, purification from the cell culture, for example, by immunoaffinity adsorption on a column (to absorb the variant by binding it to at least one remaining immune epitope). The activity of the cell lysate or purified PCA3 molecule variant is then screened in a suitable screening assay for the desired characteristic. For example, a change in the immunological character of the PCA3 molecule, such as affinity for a given antibody, is measured by a competitive type immunoassay. Changes in immunomodulation activity are measured by the appropriate assay. Modifications of such protein properties as redox or thermal stability, hydrophobicity, susceptibility to proteolytic degradation or the tendency to aggregate with carriers or into multimers are assayed by methods well known to the ordinarily skilled artisan.

A variety of methodologies known in the art can be utilized to obtain the peptide of the present invention. In one embodiment, the peptide is purified from tissues or cells which naturally produce the peptide. Alternatively, the above-described isolated nucleic acid fragments can be used to express the PCA3 protein in any organism. The samples of the present invention include cells, protein extracts or membrane extracts of cells, or biological fluids. The sample will vary based on the assay format, the

detection method and the nature of the tissues, cells or extracts used as the sample.

Any organism can be used as a source for the peptide of the invention, as long as the source organism naturally contains such a peptide. As used herein, "source organism" refers to the original organism from which the amino acid sequence of the subunit is derived, regardless of the organism the subunit is expressed in and ultimately isolated from.

One skilled in the art can readily follow known methods for isolating proteins in order to obtain the peptide free of natural contaminants. These include, but are not limited to: immunochromatography, size-exclusion chromatography, HPLC, ion-exchange chromatography, and immuno-affinity chromatography.

In a preferred embodiment, the purification procedures comprise ion-exchange chromatography and size exclusion chromatography. Any one of a large number of ion-exchange resins known in the art can be employed, including for example, monoQ, sepharose Q, macro-prepQ, AG1-X2, or HQ. Examples of suitable size exclusion resins include, but are not limited to, Superdex 200, Superose 12, and Sephacryl 200. Elution can be achieved with aqueous solutions of potassium chloride or sodium chloride at concentrations ranging from 0.01M to 2.0M.

III. A Nucleic Acid for the Specific Detection of PCA3 Nucleic Acid

In another embodiment, the present invention relates to a nucleic acid for the specific detection of the presence of PCA3 nucleic acid in a sample comprising the above-described nucleic acid molecules or at least a fragment thereof which binds under stringent conditions to PCA3 nucleic acid.

In one preferred embodiment, the present invention relates to an isolated nucleic acid consisting of 10 to 1000 nucleotides (preferably,

5

10

15

15

20

25

25

[illegible][illegible][illegible][illegible][illegible]

TABLE 2: NUCLEIC ACID PROBES (Continued)

		Size (no. of bases)	Nucleotides
	Probe 15	20	447-466 of SEQ ID NO:1
	Exon 1	120	1-120 of SEQ ID NO:6
5	Exon 2	165	121-285 of SEQ ID NO:6
	Exon 3	183	286-468 of SEQ ID NO:6
	Exon 4a	539	469-1007 of SEQ ID NO:6
	Exon 4b	1059	1008-2066 of SEQ ID NO:6
	Exon 4c	556	2067-2622 of SEQ ID NO:6
10	Exon 4d	960	2623-3582 of SEQ ID NO:6

Of course, as will be understood by the person of ordinary skill, a multitude of additional probes can be designed from the same or other region of SEQ ID NO:1 as well as from SEQ ID NO:6 and other sequences of the present invention.

The nucleic acid probe can be used to probe an appropriate chromosomal or cDNA library by usual hybridization methods to obtain another nucleic acid molecule of the present invention. A chromosomal DNA or cDNA library can be prepared from appropriate cells according to recognized methods in the art (cf. *Molecular Cloning: A Laboratory Manual, second edition*, edited by Sambrook, Fritsch, & Maniatis, Cold Spring Harbor Laboratory, 1989).

In the alternative, chemical synthesis is carried out in order to obtain nucleic acid probes having nucleotide sequences which correspond to N-terminal and C-terminal portions of the PCA3 amino acid sequence. Thus, the synthesized nucleic acid probes can be used as primers in a polymerase chain reaction (PCR) carried out in accordance with recognized PCR techniques, essentially according to *PCR Protocols, A Guide to Methods and Applications*, edited by Michael et al., Academic Press, 1990,

utilizing the appropriate chromosomal, cDNA or cell line library to obtain the fragment of the present invention.

One skilled in the art can readily design such probes based on the sequence disclosed herein using methods of computer alignment and sequence analysis known in the art (cf. *Molecular Cloning: A Laboratory Manual, second edition*, edited by Sambrook, Fritsch, & Maniatis, Cold Spring Harbor Laboratory, 1989).

The hybridization probes of the present invention can be labeled by standard labeling techniques such as with a radiolabel, enzyme label, fluorescent label, biotin-avidin label, chemiluminescence, and the like. After hybridization, the probes can be visualized using known methods.

The nucleic acid probes of the present invention include RNA, as well as DNA probes, such probes being generated using techniques known in the art.

In one embodiment of the above described method, a nucleic acid probe is immobilized on a solid support. Examples of such solid supports include, but are not limited to, plastics such as polycarbonate, complex carbohydrates such as agarose and sepharose, and acrylic resins, such as polyacrylamide and latex beads. Techniques for coupling nucleic acid probes to such solid supports are well known in the art.

The test samples suitable for nucleic acid probing methods of the present invention include, for example, cells or nucleic acid extracts of cells, or biological fluids. The sample used in the above-described methods will vary based on the assay format, the detection method and the nature of the tissues, cells or extracts to be assayed. Methods for preparing nucleic acid extracts of cells are well known in the art and can be readily adapted in order to obtain a sample which is compatible with the method utilized.

IV. A Method of Detecting The Presence of PCA3 Nucleic Acid in a Sample

In another embodiment, the present invention relates to a method of detecting the presence of PCA3 nucleic acid in a sample comprising a) contacting the sample with the above-described nucleic acid probe, under specific hybridization conditions such that hybridization occurs, and b) detecting the presence of the probe bound to the nucleic acid molecule. One skilled in the art would select the nucleic acid probe according to techniques known in the art as described above. Samples to be tested include but should not be limited to RNA or DNA samples from human tissue.

V. A Kit for Detecting the Presence of PCA3 Nucleic Acid in a Sample

In another embodiment, the present invention relates to a kit for detecting the presence of PCA3 nucleic acid in a sample comprising at least one container means having disposed therein the above-described nucleic acid probe. In a preferred embodiment, the kit further comprises other containers comprising one or more of the following: wash reagents and reagents capable of detecting the presence of bound nucleic acid probe. Examples of detection reagents include, but are not limited to radiolabelled probes, enzymatic labeled probes (horse radish peroxidase, alkaline phosphatase), and affinity labeled probes (biotin, avidin, or streptavidin).

In detail, a compartmentalized kit includes any kit in which reagents are contained in separate containers. Such containers include small glass containers, plastic containers or strips of plastic or paper. Such containers allow the efficient transfer of reagents from one compartment to another compartment such that the samples and reagents are not cross-contaminated and the agents or solutions of each container can be added in a quantitative fashion from one compartment to another. Such containers will include a container which will accept the test sample, a container which

contains the probe or primers used in the assay, containers which contain wash reagents (such as phosphate buffered saline, Tris-buffers, and the like), and containers which contain the reagents used to detect the hybridized probe, bound antibody, amplified product, or the like.

5 One skilled in the art will readily recognize that the nucleic acid probes described in the present invention can readily be incorporated into one of the established kit formats which are well known in the art.

10 **VI. DNA Constructs Comprising a PCA3 Nucleic Acid Molecule and Cells Containing These Constructs**

In another embodiment, the present invention relates to a recombinant DNA molecule comprising, 5' to 3', a promoter effective to initiate transcription in a host cell and the above-described nucleic acid molecules. In another embodiment, the present invention relates to a
15 recombinant DNA molecule comprising a vector and an above-described nucleic acid molecule.

In another embodiment, the present invention relates to a nucleic acid molecule comprising a transcriptional control region functional in a cell, a sequence complimentary to an RNA sequence encoding an amino
20 acid sequence corresponding to the above-described polypeptide, and a transcriptional termination region functional in the cell.

Preferably, the above-described molecules are isolated and/or purified DNA molecules.

In another embodiment, the present invention relates to a
25 cell or non-human organism that contains an above-described nucleic acid molecule.

In another embodiment, the peptide is purified from cells which have been altered to express the peptide.

As used herein, a cell is said to be "altered to express a desired peptide" when the cell, through genetic manipulation, is made to produce a protein which it normally does not produce or which the cell normally produces at low levels. One skilled in the art can readily adapt
5 procedures for introducing and expressing either genomic, cDNA, or synthetic sequences into either eukaryotic or prokaryotic cells.

A nucleic acid molecule, such as DNA, is said to be "capable of expressing" a polypeptide if it contains nucleotide sequences which contain transcriptional and translational regulatory information and
10 such sequences are "operably linked" to nucleotide sequences which encode the polypeptide. An operable linkage is a linkage in which the regulatory DNA sequences and the DNA sequence sought to be expressed are connected in such a way as to permit gene sequence expression. The precise nature of the regulatory regions needed for gene sequence expression can vary from
15 organism to organism, but shall in general include a promoter region which, in prokaryotes, contains both the promoter (which directs the initiation of RNA transcription) as well as the DNA sequences which, when transcribed into RNA, will signal synthesis initiation. Such regions will normally include those 5'-non-coding sequences involved with initiation of transcription and
20 translation, such as the TATA box, capping sequence, CAAT sequence, and the like.

If desired, the non-coding region 3' to the PCA3 coding sequence can be obtained by the above-described methods. This region can be retained for its transcriptional termination regulatory sequences, such as
25 termination and polyadenylation. Thus, by retaining the 3'-region naturally contiguous to the DNA sequence encoding a PCA3 gene, the transcriptional termination signals can be provided. Where the transcriptional termination signals are not satisfactorily functional in the expression host cell, then a 3' region functional in the host cell can be substituted.

Two DNA sequences (such as a promoter region sequence and a PCA3 coding sequence) are said to be operably linked if the nature of the linkage between the two DNA sequences does not (1) result in the introduction of a frame-shift mutation, (2) interfere with the ability of the promoter region sequence to direct the transcription of a PCA3 coding sequence, or (3) interfere with the ability of the PCA3 coding sequence to be transcribed by the promoter region sequence. Thus, a promoter region would be operably linked to a DNA sequence if the promoter were capable of effecting transcription of that DNA sequence.

The present invention encompasses the expression of the PCA3 coding sequence (or a functional derivative thereof) in either prokaryotic or eukaryotic cells. Prokaryotic hosts are, generally, the most efficient and convenient for the production of recombinant proteins and, therefore, are preferred for the expression of the PCA3 coding sequence.

Prokaryotes most frequently are represented by various strains of *E. coli*. However, other microbial strains can also be used, including other bacterial strains. In prokaryotic systems, plasmid vectors that contain replication sites and control sequences derived from a species compatible with the host can be used. Examples of suitable plasmid vectors include pBR322, pUC18, pUC19, pUC118, pUC119 and the like; suitable phage or bacteriophage vectors include λ gt10, λ gt11 and the like; and suitable virus vectors include pMAM-neo, pKRC and the like. Preferably, the selected vector of the present invention has the capacity to replicate in the selected host cell.

Recognized prokaryotic hosts include bacteria such as *E. coli*, *Bacillus*, *Streptomyces*, *Pseudomonas*, *Salmonella*, *Serratia*, and the like. However, under such conditions, the peptide will not be glycosylated. The prokaryotic host must be compatible with the replicon and control sequences in the expression plasmid.

To express PCA3 in a prokaryotic cell, it is necessary to operably link the PCA3 coding sequence to a functional prokaryotic promoter. Such promoters can be either constitutive or, more preferably, regulatable (i.e., inducible or derepressible). Examples of constitutive promoters include the *int* promoter of bacteriophage λ , the *bla* promoter of the β -lactamase gene sequence of pBR322, and the CAT promoter of the chloramphenicol acetyl transferase gene sequence of pBR325, and the like. Examples of inducible prokaryotic promoters include the major right and left promoters of bacteriophage λ (P_L and P_R), the *trp*, *recA*, *lacZ*, *lacI*, and *gal* promoters of *E. coli*, the α -amylase (Ulmanen *et al.*, *J. Bacteriol.* 162:176-182 (1985)) and the ζ -28-specific promoters of *B. subtilis* (Gilman *et al.*, *Gene sequence* 32:11-20 (1984)), the promoters of the bacteriophages of *Bacillus* (Gryczan, In: *The Molecular Biology of the Bacilli*, Academic Press, Inc., NY (1982)), and *Streptomyces* promoters (Ward *et al.*, *Mol. Gen. Genet.* 203:468-478 (1986)). Prokaryotic promoters are reviewed by Glick (*J. Ind. Microbiol.* 1:277-282 (1987)); Cenatiempo (*Biochimie* 68:505-516 (1986)); and Gottesman (*Ann. Rev. Genet.* 18:415-442 (1984)).

Proper expression in a prokaryotic cell also requires the presence of a ribosome binding site upstream of the gene sequence-encoding sequence. Such ribosome binding sites are disclosed, for example, by Gold *et al.* (*Ann. Rev. Microbiol.* 35:365-404 (1981)).

The selection of control sequences, expression vectors, transformation methods, and the like, are dependent on the type of host cell used to express the gene. As used herein, "cell", "cell line", and "cell culture" can be used interchangeably and all such designations include progeny. Thus, the words "transformants" or "transformed cells" include the primary subject cell and cultures derived therefrom, without regard to the number of transfers. It is also understood that all progeny can not be precisely identical in DNA content, due to deliberate or inadvertent mutations. However, as

defined, mutant progeny have the same functionality as that of the originally transformed cell.

Host cells which can be used in the expression systems of the present invention are not strictly limited, provided that they are suitable for use in the expression of the PCA3 peptide of interest. Suitable hosts include eukaryotic cells.

Preferred eukaryotic hosts include, for example, yeast, fungi, insect cells, mammalian cells either *in vivo*, or in tissue culture. Preferred mammalian cells include HeLa cells, cells of fibroblast origin such as VERO or CHO-K1, or cells of lymphoid origin and their derivatives.

In addition, plant cells are also available as hosts, and control sequences compatible with plant cells are available, such as the cauliflower mosaic virus 35S and 19S, and nopaline synthase promoter and polyadenylation signal sequences.

Another preferred host is an insect cell, for example *Drosophila* larvae. Using insect cells as hosts, the *Drosophila* alcohol dehydrogenase promoter can be used, Rubin, *Science* 240:1453-1459 (1988). Alternatively, baculovirus vectors can be engineered to express large amounts of PCA3 in insect cells (Jasny, *Science* 238:1653 (1987); Miller *et al.*, In: *Genetic Engineering* (1986), Setlow, J.K., *et al.*, eds., Plenum, Vol. 8, pp. 277-297).

Different host cells have characteristic and specific mechanisms for the translational and post-translational processing and modification (e.g., glycosylation, cleavage) of proteins. Appropriate cell lines or host systems can be chosen to ensure the desired modification and processing of the foreign protein expressed.

Any of a series of yeast gene sequence expression systems can be utilized which incorporate promoter and termination elements from the actively expressed gene sequences coding for glycolytic enzymes.

These enzymes are produced in large quantities when yeast are grown in mediums rich in glucose. Known glycolytic gene sequences can also provide very efficient transcriptional control signals.

Yeast provides substantial advantages in that it can also
5 carry out post-translational peptide modifications. A number of recombinant DNA strategies exist which utilize strong promoter sequences and high copy number of plasmids which can be utilized for production of the desired proteins in yeast. Yeast recognizes leader sequences on cloned mammalian gene sequence products and secretes peptides bearing leader sequences
10 (i.e., pre-peptides). For a mammalian host, several possible vector systems are available for the expression of PCA3.

A wide variety of transcriptional and translational regulatory sequences can be employed, depending upon the nature of the host. The transcriptional and translational regulatory signals can be derived from viral
15 sources, such as adenovirus, bovine papilloma virus, simian virus, or the like, where the regulatory signals are associated with a particular gene sequence which has a high level of expression. Alternatively, promoters from mammalian expression products, such as actin, collagen, myosin, and the like, can be employed. Transcriptional initiation regulatory signals can be
20 selected which allow for repression or activation, so that expression of the gene sequences can be modulated. Of interest are regulatory signals which are temperature-sensitive so that by varying the temperature, expression can be repressed or initiated, or are subject to chemical (such as metabolite) regulation.

25 As discussed above, expression of PCA3 in eukaryotic hosts requires the use of eukaryotic regulatory regions. Such regions will, in general, include a promoter region sufficient to direct the initiation of RNA synthesis. Preferred eukaryotic promoters include, for example, the promoter of the mouse metallothionein I gene sequence (Hamer *et al.*, *J. Mol. Appl.*

Gen. 1:273-288 (1982)); the TK promoter of Herpes virus (McKnight, *Cell* 31:355-365 (1982)); the SV40 early promoter (Benoist *et al.*, *Nature (London)* 290:304-310 (1981)); the yeast *gal4* gene sequence promoter (Johnston *et al.*, *Proc. Natl. Acad. Sci. (USA)* 79:6971-6975 (1982); Silver *et al.*, *Proc. Natl. Acad. Sci. (USA)* 81:5951-5955 (1984)) and the CMV immediate-early gene promoter (Thomsen *et al.*, *Proc. Natl. Acad. Sci (USA)* 81:659-663 (1984)).

As is widely known, translation of eukaryotic mRNA is initiated at the codon which encodes the first methionine. For this reason, it is preferable to ensure that the linkage between a eukaryotic promoter and a PCA3 coding sequence does not contain any intervening codons which are capable of encoding a methionine (i.e., AUG). The presence of such codons results either in a formation of a fusion protein (if the AUG codon is in the same reading frame as the PCA3 coding sequence) or a frame-shift mutation (if the AUG codon is not in the same reading frame as the PCA3 coding sequence).

A PCA3 nucleic acid molecule and an operably linked promoter can be introduced into a recipient prokaryotic or eukaryotic cell either as a non-replicating DNA (or RNA) molecule, which can either be a linear molecule or, more preferably, a closed covalent circular molecule. Since such molecules are incapable of autonomous replication, the expression of the gene can occur through the transient expression of the introduced sequence. Alternatively, permanent expression can occur through the integration of the introduced DNA sequence into the host chromosome.

In one embodiment, a vector is employed which is capable of integrating the desired gene sequences into the host cell chromosome. Cells which have stably integrated the introduced DNA into their chromosomes can be selected by also introducing one or more markers which allow for selection of host cells which contain the expression vector.

09402713-061300

The marker can provide for prototrophy to an auxotrophic host, biocide resistance, e.g., antibiotics, or heavy metals, such as copper, or the like. The selectable marker gene sequence can either be directly linked to the DNA gene sequences to be expressed, or introduced into the same cell by co-

5 transfection. Additional elements can also be needed for optimal synthesis of single chain binding protein mRNA. These elements can include splice signals, as well as transcription promoters, enhancer signal sequences, and termination signals. cDNA expression vectors incorporating such elements include those described by Okayama, *Molec. Cell. Biol.* 3:280 (1983).

10 In a preferred embodiment, the introduced nucleic acid molecule will be incorporated into a plasmid or viral vector capable of autonomous replication in the recipient host. Any of a wide variety of vectors can be employed for this purpose. Factors of importance in selecting a particular plasmid or viral vector include: the ease with which recipient cells

15 that contain the vector can be recognized and selected from those recipient cells which do not contain the vector; the number of copies of the vector which are desired in a particular host; and whether it is desirable to be able to "shuttle" the vector between host cells of different species. Preferred prokaryotic vectors include plasmids such as those capable of replication in

20 *E. coli* (such as, for example, pBR322, ColE1, pSC101, pACYC 184, π VX. Such plasmids are, for example, disclosed by Sambrook (cf. *Molecular Cloning: A Laboratory Manual*, second edition, edited by Sambrook, Fritsch, & Maniatis, Cold Spring Harbor Laboratory, 1989). *Bacillus* plasmids include pC194, pC221, pT127, and the like. Such plasmids are disclosed by Gryczan

25 (In: *The Molecular Biology of the Bacilli*, Academic Press, NY (1982), pp. 307-329). Suitable *Streptomyces* plasmids include pIJ101 (Kendall *et al.*, *J. Bacteriol.* 169:4177-4183 (1987)), and streptomyces bacteriophages such as ϕ C31 (Chater *et al.*, In: *Sixth International Symposium on Actinomycetales Biology*, Akademiai Kiado, Budapest, Hungary (1986), pp. 45-54).

Pseudomonas plasmids are reviewed by John *et al.* (*Rev. Infect. Dis.* 8:693-704 (1986)), and Izaki (*Jpn. J. Bacteriol.* 33:729-742 (1978)).

Preferred eukaryotic plasmids include, for example, BPV, vaccinia, SV40, 2-micron circle, and the like, or their derivatives. Such plasmids are well known in the art (Botstein *et al.*, *Miami Wntr. Symp.* 19:265-274 (1982); Broach, In: *The Molecular Biology of the Yeast Saccharomyces: Life Cycle and Inheritance*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, p. 445-470 (1981); Broach, *Cell* 28:203-204 (1982); Bollon *et al.*, *J. Clin. Hematol. Oncol.* 10:39-48 (1980); Maniatis, In: *Cell Biology: A Comprehensive Treatise*, Vol. 3, Gene Sequence Expression, Academic Press, NY, pp. 563-608 (1980)).

Once the vector or nucleic acid molecule containing the construct(s) has been prepared for expression, the DNA construct(s) can be introduced into an appropriate host cell by any of a variety of suitable means, i.e., transformation, transfection, conjugation, protoplast fusion, electroporation, particle gun technology, calcium phosphate-precipitation, direct microinjection, and the like. After the introduction of the vector, recipient cells are grown in a selective medium, which selects for the growth of vector-containing cells. Expression of the cloned gene molecule(s) results in the production of PCA3. This can take place in the transformed cells as such, or following the induction of these cells to differentiate (for example, by administration of bromodeoxyuracil to neuroblastoma cells or the like).

VII. An Antibody Having Binding Affinity to a PCA3 Polypeptide and a Hybridoma Containing the Antibody

In another embodiment, the present invention relates to an antibody having binding affinity specifically to a PCA3 polypeptide as described above or specifically to a PCA3 polypeptide binding fragment thereof. An antibody binds specifically to a PCA3 polypeptide or binding

fragment thereof if it does not bind to non-PCA3 polypeptides. Those which bind selectively to PCA3 would be chosen for use in methods which could include, but should not be limited to, the analysis of altered PCA3 expression in tissue containing PCA3.

5 The PCA3 proteins of the present invention can be used in a variety of procedures and methods, such as for the generation of antibodies, for use in identifying pharmaceutical compositions, and for studying DNA/protein interaction.

10 The PCA3 peptide of the present invention can be used to produce antibodies or hybridomas. One skilled in the art will recognize that if an antibody is desired, such a peptide would be generated as described herein and used as an immunogen.

15 The antibodies of the present invention include monoclonal and polyclonal antibodies, as well as fragments of these antibodies. The invention further includes single chain antibodies. Antibody fragments which contain the idiotype of the molecule can be generated by known techniques. For example, such fragments include but are not limited to: the F(ab')₂ fragment; the Fab' fragments, Fab fragments, and Fv fragments.

20 Of special interest to the present invention are antibodies to PCA3 which are produced in humans, or are "humanized" (i.e. non-immunogenic in a human) by recombinant or other technology. Humanized antibodies can be produced, for example by replacing an immunogenic portion of an antibody with a corresponding, but non-immunogenic portion (i.e. chimeric antibodies) (Robinson, R.R. *et al.*, International Patent Publication PCT/US86/02269; Akira, K. *et al.*, European Patent Application 184,187; Taniguchi, M., European Patent Application 171,496; Morrison, S.L. *et al.*, European Patent Application 173,494; Neuberger, M.S. *et al.*, PCT Application WO 86/01533; Cabilly, S. *et al.*, European Patent Application 125,023; Better, M. *et al.*, *Science* 240:1041-1043 (1988); Liu, A.Y. *et al.*,

25

Proc. Natl. Acad. Sci. USA 84:3439-3443 (1987); Liu, A.Y. et al., J. Immunol. 139:3521-3526 (1987); Sun, L.K. et al., Proc. Natl. Acad. Sci. USA 84:214-218 (1987); Nishimura, Y. et al., Canc. Res. 47:999-1005 (1987); Wood, C.R. et al., Nature 314:446-449 (1985)); Shaw et al., J. Natl. Cancer Inst. 80:1553-1559 (1988). General reviews of "humanized" chimeric antibodies are provided by Morrison, S.L. (Science, 229:1202-1207 (1985)) and by Oi, V.T. et al., BioTechniques 4:214 (1986)). Suitable "humanized" antibodies can be alternatively produced by CDR or CEA substitution (Jones, P.T. et al., Nature 321:552-525 (1986); Verhoeyan et al., Science 239:1534 (1988); Beidler, C.B. et al., J. Immunol. 141:4053-4060 (1988)).

In another embodiment, the present invention relates to a hybridoma which produces the above-described monoclonal antibody. A hybridoma is an immortalized cell line which is capable of secreting a specific monoclonal antibody.

In general, techniques for preparing monoclonal antibodies and hybridomas are well known in the art (Campbell, "Monoclonal Antibody Technology: Laboratory Techniques in Biochemistry and Molecular Biology," Elsevier Science Publishers, Amsterdam, The Netherlands (1984); St. Groth et al., J. Immunol. Methods 35:1-21 (1980)).

Any animal (mouse, rabbit, and the like) which is known to produce antibodies can be immunized with the selected polypeptide. Methods for immunization are well known in the art. Such methods include subcutaneous or interperitoneal injection of the polypeptide. One skilled in the art will recognize that the amount of polypeptide used for immunization will vary based on the animal which is immunized, the antigenicity of the polypeptide and the site of injection.

The polypeptide can be modified or administered in an adjuvant in order to increase the peptide antigenicity. Methods of increasing the antigenicity of a polypeptide are well known in the art. Such procedures

include coupling the antigen with a heterologous protein (such as globulin or β -galactosidase) or through the inclusion of an adjuvant during immunization.

For monoclonal antibodies, spleen cells from the immunized animals are removed, fused with myeloma cells, and allowed to become monoclonal antibody producing hybridoma cells.

Any one of a number of methods well known in the art can be used to identify the hybridoma cell which produces an antibody with the desired characteristics. These include screening the hybridomas with an ELISA assay, western blot analysis, or radioimmunoassay (Lutz *et al.*, *Exp. Cell Res.* 175:109-124 (1988)).

Hybridomas secreting the desired antibodies are cloned and the class and subclass is determined using procedures known in the art (Campbell, *Monoclonal Antibody Technology: Laboratory Techniques in Biochemistry and Molecular Biology*, *supra* (1984)).

For polyclonal antibodies, antibody containing antisera is isolated from the immunized animal and is screened for the presence of antibodies with the desired specificity using one of the above-described procedures.

In another embodiment of the present invention, the above-described antibodies are detectably labeled. Antibodies can be detectably labeled through the use of radioisotopes, affinity labels (such as biotin, avidin, and the like), enzymatic labels (such as horse radish peroxidase, alkaline phosphatase, and the like) fluorescent labels (such as FITC or rhodamine, and the like), paramagnetic atoms, and the like. Procedures for accomplishing such labeling are well-known in the art, for example, see (Sternberger *et al.*, *J. Histochem. Cytochem.* 18:315 (1970); Bayer *et al.*, *Meth. Enzym.* 62:308 (1979); Engval *et al.*, *Immunol.* 109:129 (1972); Goding, *J. Immunol. Meth.* 13:215 (1976)). The labeled antibodies of the

present invention can be used for *in vitro*, *in vivo*, and *in situ* assays to identify cells or tissues which express a specific peptide.

In another embodiment of the present invention the above-described antibodies are immobilized on a solid support. Examples of such solid supports include plastics such as polycarbonate, complex carbohydrates such as agarose and sepharose, acrylic resins and such as polyacrylamide and latex beads. Techniques for coupling antibodies to such solid supports are well known in the art (Weir *et al.*, "*Handbook of Experimental Immunology*" 4th Ed., Blackwell Scientific Publications, Oxford, England, Chapter 10 (1986); Jacoby *et al.*, *Meth. Enzym.* 34 Academic Press, N.Y. (1974)). The immobilized antibodies of the present invention can be used for *in vitro*, *in vivo*, and *in situ* assays as well as in immunochromatography.

Furthermore, one skilled in the art can readily adapt currently available procedures, as well as the techniques, methods and kits disclosed above with regard to antibodies, to generate peptides capable of binding to a specific peptide sequence in order to generate rationally designed antipeptide peptides, for example see Hurby *et al.*, "Application of Synthetic Peptides: Antisense Peptides", In *Synthetic Peptides, A User's Guide*, W.H. Freeman, NY, pp. 289-307 (1992), and Kaspczak *et al.*, *Biochemistry* 28:9230-8 (1989).

Anti-peptide peptides can be generated in one of two fashions. First, the anti-peptide peptides can be generated by replacing the basic amino acid residues found in the PCA3 peptide sequence with acidic residues, while maintaining hydrophobic and uncharged polar groups. For example, lysine, arginine, and/or histidine residues are replaced with aspartic acid or glutamic acid and glutamic acid residues are replaced by lysine, arginine or histidine.

VIII. A Method of Detecting a PCA3 Polypeptide or Antibody in a Sample

In another embodiment, the present invention relates to a method of detecting a PCA3 polypeptide in a sample, comprising: a) contacting the sample with an above-described antibody (or protein), under conditions such that immunocomplexes form, and b) detecting the presence of the antibody bound to the polypeptide. In detail, the methods comprise incubating a test sample with one or more of the antibodies of the present invention and assaying whether the antibody binds to the test sample. Altered levels of PCA3 in a sample as compared to normal levels can indicate a specific disease (ex. prostate cancer).

In a further embodiment, the present invention relates to a method of detecting a PCA3 antibody in a sample, comprising: a) contacting the sample with an above-described PCA3 protein, under conditions such that immunocomplexes form, and b) detecting the presence of the protein bound to the antibody or antibody bound to the protein. In detail, the methods comprise incubating a test sample with one or more of the proteins of the present invention and assaying whether the antibody binds to the test sample.

Conditions for incubating an antibody with a test sample vary. Incubation conditions depend on the format employed in the assay, the detection methods employed, and the type and nature of the antibody used in the assay. One skilled in the art will recognize that any one of the commonly available immunological assay formats (such as radioimmunoassays, enzyme-linked immunosorbent assays, diffusion based Ouchterlony, or rocket immunofluorescent assays) can readily be adapted to employ the antibodies of the present invention. Examples of such assays can be found in Chard, *An Introduction to Radioimmunoassay and Related Techniques*, Elsevier Science Publishers, Amsterdam, The Netherlands

(1986); Bullock *et al.*, *Techniques in Immunocytochemistry*, Academic Press, Orlando, FL Vol. 1 (1982), Vol. 2 (1983), Vol. 3 (1985); Tijssen, *Practice and Theory of Enzyme Immunoassays: Laboratory Techniques in Biochemistry and Molecular Biology*, Elsevier Science Publishers, Amsterdam, The Netherlands (1985).

The immunological assay test samples of the present invention include cells, protein or membrane extracts of cells, or biological fluids such as blood, serum, plasma, or urine. The test sample used in the above-described method will vary based on the assay format, nature of the detection method and the tissues, cells or extracts used as the sample to be assayed. Methods for preparing protein extracts or membrane extracts of cells are well known in the art and can be readily be adapted in order to obtain a sample which is capable with the system utilized.

IX. A Diagnostic Kit Comprising PCA3 Protein or Antibody

In another embodiment of the present invention, a kit is provided which contains all the necessary reagents to carry out the previously described methods of detection.

The kit can comprise: i) a first container means containing an above-described antibody, and ii) second container means containing a conjugate comprising a binding partner of the antibody and a label.

The kit can comprise: i) a first container means containing an above-described protein, and preferably, ii) second container means containing a conjugate comprising a binding partner of the protein and a label. More specifically, a diagnostic kit comprises PCA3 protein as described above, to detect antibodies in the serum of potentially infected animals or humans.

In another preferred embodiment, the kit further comprises one or more other containers comprising one or more of the following: wash

reagents and reagents capable of detecting the presence of bound antibodies. Examples of detection reagents include, but are not limited to, labeled secondary antibodies, or in the alternative, if the primary antibody is labeled, the chromophoric, enzymatic, or antibody binding reagents which are capable of reacting with the labeled antibody. The compartmentalized kit can be as described above for nucleic acid probe kits.

One skilled in the art will readily recognize that the antibodies described in the present invention can readily be incorporated into one of the established kit formats which are well known in the art.

X. Diagnostic Screening

It is to be understood that although the following discussion is specifically directed to human patients, the teachings are also applicable to any animal that expresses PCA3.

The diagnostic and screening methods of the invention are especially useful for a patient suspected of being at risk for developing a disease associated with an altered expression level of PCA3 based on family history, or a patient in which it is desired to diagnose a PCA3-related disease (ex. prostate cancer).

According to the invention, presymptomatic screening of an individual in need of such screening is now possible using DNA encoding the PCA3 protein or the PCA3 gene of the invention or fragments thereof. The screening method of the invention allows a presymptomatic diagnosis, including prenatal diagnosis, of the presence of a missing or aberrant PCA3 gene in individuals, and thus an opinion concerning the likelihood that such individual would develop or has developed a PCA3-associated disease. This is especially valuable for the identification of carriers of altered or missing PCA3 genes, for example, from individuals with a family history of a PCA3-

associated disease. Early diagnosis is also desired to maximize appropriate timely intervention.

In one preferred embodiment of the method of screening, a tissue sample would be taken from such individual, and screened for (1) the presence of the "normal" PCA3 gene; (2) the presence of PCA3 mRNA and/or (3) the presence of PCA3 protein. The normal human gene can be characterized based upon, for example, detection of restriction digestion patterns in "normal" versus the patient's DNA, including RFLP analysis, using DNA probes prepared against the PCA3 sequence (or a functional fragment thereof) taught in the invention. Similarly, PCA3 mRNA can be characterized and compared to normal PCA3 mRNA (a) levels and/or (b) size as found in a human population not at risk of developing PCA3-associated disease using similar probes. Lastly, PCA3 protein can be (a) detected and/or (b) quantitated using a biological assay for PCA3 activity or using an immunological assay and PCA3 antibodies. When assaying PCA3 protein, the immunological assay is preferred for its speed. An (1) aberrant PCA3 DNA size pattern, and/or (2) aberrant PCA3 mRNA sizes or levels and/or (3) aberrant PCA3 protein levels would indicate that the patient is at risk for developing a PCA3-associated disease.

More specifically, a method of diagnosing the presence or predisposition to develop prostate cancer in a patient is provided herein.

The screening and diagnostic methods of the invention do not require that the entire PCA3 DNA coding sequence be used for the probe. Rather, it is only necessary to use a fragment or length of nucleic acid that is sufficient to detect the presence of the PCA3 gene in a DNA preparation from a normal or affected individual, the absence of such gene, or an altered physical property of such gene (such as a change in electrophoretic migration pattern). Preferably, any of the probes as described above are used.

Prenatal diagnosis can be performed when desired, using any known method to obtain fetal cells, including amniocentesis, chorionic villous sampling (CVS), and fetoscopy. Prenatal chromosome analysis can be used to determine if the portion of the chromosome possessing the normal PCA3 gene is present in a heterozygous state.

XI. Therapeutic Treatments

A. Therapeutic Nucleic Acids

A therapeutic nucleic acid as a therapeutic agent can have, but is not limited to, at least one of the following therapeutic effects on a target cell: inhibiting transcription of a DNA sequence; inhibiting translation of an RNA sequence; inhibiting reverse transcription of an RNA or DNA sequence; inhibiting a post-translational modification of a protein; inducing transcription of a DNA sequence; inducing translation of an RNA sequence; inducing reverse transcription of an RNA or DNA sequence; inducing a post-translational modification of a protein; transcription of the nucleic acid as an RNA; translation of the nucleic acid as a protein or enzyme; and incorporating the nucleic acid into a chromosome of a target cell for constitutive or transient expression of the therapeutic nucleic acid.

Therapeutic effects of therapeutic nucleic acids can include, but are not limited to: turning off a defective gene or processing the expression thereof, such as antisense RNA or DNA; inhibiting viral replication or synthesis; gene therapy as expressing a heterologous nucleic acid encoding a therapeutic protein or correcting a defective protein; modifying a defective or underexpression of an RNA such as an hnRNA, an mRNA, a tRNA, or an rRNA; encoding a drug or prodrug, or an enzyme that generates a compound as a drug or prodrug in pathological or normal cells expressing the chimeric receptor; and any other known therapeutic effects.

In the method of treating a PCA3-associated disease (preferably, prostate cancer) in a patient in need of such treatment, a PCA3

gene which is not indicative of a disease state can be provided to the cells of such patient in a manner and amount that permits the expression of the PCA3 protein provided by such gene, for a time and in a quantity sufficient to treat such patient. Preferably, gene replacement ("knock out") technology is used that would replace the disease causing PCA3 gene with a PCA3 gene which does not cause disease (specifically, prostate cancer).

Included as well in the invention are pharmaceutical compositions comprising an effective amount of at least one PCA3 antisense oligonucleotide, in combination with a pharmaceutically acceptable carrier.

Such antisense oligos include, but are not limited to, at least one nucleotide sequence of 12-500 bases in length which is complementary to PCA3 exons 1, 2, 3, 4a-4d; a DNA sequence of SEQ ID NO:1, 3, 4, or 6; or a DNA sequence encoding at least 4 amino acids of SEQ ID NO:2 or SEQ ID NO:7.

Alternatively, the PCA3 nucleic acid can be combined with a lipophilic carrier such as any one of a number of sterols including cholesterol, cholate and deoxycholic acid. A preferred sterol is cholesterol.

The PCA3 gene therapy nucleic acids and the pharmaceutical compositions of the invention can be administered by any means that achieve their intended purpose. For example, administration can be by parenteral, subcutaneous, intravenous, intramuscular, intra-peritoneal, or transdermal routes. The dosage administered will be dependent upon the age, health, and weight of the recipient, kind of concurrent treatment, if any, frequency of treatment, and the nature of the effect desired.

Compositions within the scope of this invention include all compositions wherein the PCA3 antisense oligonucleotide is contained in an amount effective to achieve decreased expression of at least one PCA3 gene. While individual needs vary, determination of optimal ranges of effective amounts of each component is within the skill of the art. Typically, the PCA3 nucleic acid can be administered to mammals, e.g. humans, at a

dose of 0.005 to 1 mg/kg/day, or an equivalent amount of the pharmaceutically acceptable salt thereof, per day of the body weight of the mammal being treated.

Suitable formulations for parenteral administration include aqueous solutions of the PCA3 nucleic acid in water-soluble form, for example, water-soluble salts. In addition, suspensions of the active compounds as appropriate oily injection suspensions can be administered. Suitable lipophilic solvents or vehicles include fatty oils, for example, sesame oil, or synthetic fatty acid esters, for example, ethyl oleate or triglycerides. Aqueous injection suspensions can contain substances which increase the viscosity of the suspension include, for example, sodium carboxymethyl cellulose, sorbitol, and/or dextran. Optionally, the suspension can also contain stabilizers.

Many vector systems are known in the art to provide such delivery to human patients in need of a gene or protein missing from the cell. For example, retrovirus systems can be used, especially modified retrovirus systems and especially herpes simplex virus systems (Gage *et al.*, U.S. Patent No. 5,082,670). Such methods are provided for, in, for example, the teachings of Breakefield, X.A. *et al.*, *The New Biologist* 3:203-218 (1991); Huang, Q. *et al.*, *Experimental Neurology* 115:303-316 (1992); WO93/03743; WO90/0944; Taylor, WO 92/06693; Mulligan, R.C., *Science* 260:926-932 (1993); and Brown *et al.*, "Retroviral Vectors," in *DNA Cloning: A Practical Approach*, Volume 3, IRL Press, Washington, D.C. (1987). Delivery of a DNA sequence encoding a normally expressed PCA3 protein will effectively replace the PCA3 gene responsible for the disease state (ex. prostate cancer).

The means by which the vector carrying the nucleic acid can be introduced into the cell include but is not limited to, microinjection, electroporation, transduction, or transfection using DEAE-Dextran,

lipofection, calcium phosphate or other procedures known to one skilled in the art (*Molecular Cloning, A Laboratory Manual*, Sambrook *et al.*, eds., Cold Spring Harbor Press, Plainview, New York (1989)).

5 In another embodiment of this invention, a normal PCA3 gene is expressed as a recombinant gene in a cell, so that the cells can be transplanted into a mammal, preferably a human in need of gene therapy. To provide gene therapy to an individual, a genetic sequence which encodes for all or part of the PCA3 gene is inserted into a vector and introduced into a host cell.

10 Further gene therapy methods which can be used to transfer nucleic acid to a patient are set forth in Chatterjee and Wong, *Current Topics in Microbiol. and Immuno.*, 218: 61-73 (1996); Zhang, *J. Mol. Med.* 74:191-204 (1996); Schmidt-Wolf and Schmidt-Wolf, *J. of Hematotherapy* 4:551-561 (1995); Shaughnessy *et al.*, *Seminars in Oncology* 15 23(1): 159-171 (1996); and Dunbar *Annu. Rev. Med.* 47:11-20 (1996).

Specificity for gene expression in prostate cancer cells can be conferred by using appropriate cell-specific regulatory sequences, such as cell-specific enhancers and promoters.

20 Thus, gene therapy can be used to alleviate PCA3 related pathology by inhibiting the inappropriate expression of a particular form of PCA3. Moreover, gene therapy can be used to alleviate such pathologies by providing the appropriate expression level of a particular form of PCA3. In this case, particular PCA3 nucleic acid sequences can be coded by DNA or RNA constructs which are administered in the form of viruses, as described 25 above.

B. Antagonists and Agonists of PCA3

The ability of antagonists and agonists of PCA3 to interfere or enhance the activity of PCA3 can be evaluated with cells containing PCA3.

An assay for PCA3 activity in cells can be used to determine the functionality of the PCA3 protein in the presence of an agent which may act as antagonist or agonist, and thus, agents that interfere or enhance the activity of PCA3 are identified.

5 The agents screened in the assays can be, but are not limited to, antibodies, peptides, carbohydrates, vitamin derivatives, or other pharmaceutical agents. These agents can be selected and screened 1) at random, 2) by a rational selection or 3) by design using for example, protein or ligand modeling techniques (preferably, computer modeling).

10 For random screening, agents such as antibodies, peptides, carbohydrates, pharmaceutical agents and the like are selected at random and are assayed for their ability to bind to or stimulate/block the activity of the PCA3 protein.

15 Alternatively, agents may be rationally selected or designed. As used herein, an agent is said to be "rationally selected or designed" when the agent is chosen based on the configuration of the PCA3 protein.

20 In one embodiment, the present invention relates to a method of screening for an antagonist or agonist which stimulates or blocks the activity of PCA3 comprising:

(a) incubating a cell expressing PCA3 with an agent to be tested; and

(b) assaying the cell for the activity of the PCA3 protein by measuring the agents effect on ATP binding of PCA3.

25 Any cell may be used in the above assay so long as it expresses a functional form of PCA3 and the PCA3 activity can be measured. The preferred expression cells are eukaryotic cells or organisms. Such cells can be modified to contain DNA sequences encoding PCA3 using

routine procedures known in the art. Alternatively, one skilled in the art can introduce mRNA encoding the PCA3 protein directly into the cell.

Using PCA3 ligands (ligands including antagonists and agonists as described above) the present invention further provides a method for modulating the activity of the PCA3 protein in a cell. In general, ligands (antagonists and agonists) which have been identified to block or stimulate the activity of PCA3 can be formulated so that the ligand can be contacted with a cell expressing a PCA3 protein *in vivo*. The contacting of such a cell with such a ligand results in the *in vivo* modulation of the activity of the PCA3 proteins. So long as a formulation barrier or toxicity barrier does not exist, ligands identified in the assays described above will be effective for *in vivo* use.

In another embodiment, the present invention relates to a method of administering PCA3 or a PCA3 ligand (including PCA3 antagonists and agonists) to an animal (preferably, a mammal (specifically, a human)) in an amount sufficient to effect an altered level of PCA3 in the animal. The administered PCA3 or PCA3 ligand could specifically effect PCA3 associated functions. Further, since PCA3 is expressed in prostatic cancer cells, administration of PCA3 or PCA3 ligand could be used to alter PCA3 levels in such cells.

One skilled in the art will appreciate that the amounts to be administered for any particular treatment protocol can readily be determined. The dosage should not be so large as to cause adverse side effects, such as unwanted cross-reactions, anaphylactic reactions, and the like. Generally, the dosage will vary with the age, condition, sex and extent of disease in the patient, counter indications, if any, and other such variables, to be adjusted by the individual physician. Dosage can vary from .001 mg/kg to 50 mg/kg of PCA3 or PCA3 ligand, in one or more administrations daily, for one or several days. PCA3 or PCA3 ligand can be administered parenterally by injection or

by gradual perfusion over time. It can be administered intravenously, intraperitoneally, intramuscularly, or subcutaneously.

Preparations for parenteral administration include sterile or aqueous or non-aqueous solutions, suspensions, and emulsions. Examples of non-aqueous solvents are propylene glycol, polyethylene glycol, vegetable oils such as olive oil, and injectable organic esters such as ethyl oleate. Aqueous carriers include water, alcoholic/aqueous solutions, emulsions or suspensions, including saline and buffered media. Parenteral vehicles include sodium chloride solution, Ringer's dextrose and sodium chloride, lactated Ringer's, or fixed oils. Intravenous vehicles include fluid and nutrient replenishers, electrolyte replenishers, such as those based on Ringer's dextrose, and the like. Preservatives and other additives can also be present, such as, for example, antimicrobials, antioxidants, chelating agents, inert gases and the like. See, generally, *Remington's Pharmaceutical Science*, 16th Ed., Mack Eds. (1980).

In another embodiment, the present invention relates to a pharmaceutical composition comprising PCA3 or PCA3 ligand in an amount sufficient to alter PCA3 associated activity, and a pharmaceutically acceptable diluent, carrier, or excipient. Appropriate concentrations and dosage unit sizes can be readily determined by one skilled in the art as described above (See, for example, *Remington's Pharmaceutical Sciences* (16th ed., Osol, A., Ed., Mack, Easton PA (1980) and WO 91/19008).

C. Immunotherapy

The present invention provides the above-described PCA3 antibodies (preferably, PCA3 murine antibodies and chimeric PCA3 murine-human antibodies, and fragments and regions thereof) which inhibit or neutralize PCA3 biological activity *in vivo* and are specific for PCA3. These antibodies can be used for therapeutic purposes in subjects having

pathologies or conditions associated with the presence of aberrant PCA3 expression. Antibodies, and fragments, regions and derivatives thereof, of the present invention preferably contain at least one region which recognizes an epitope of PCA3 which has inhibiting and/or neutralizing biological activity *in vivo*.

Treatment comprises parenterally administering a single or multiple doses of the antibody, fragment or derivative. Preferred for human pharmaceutical use are high affinity potent PCA3-inhibiting and/or neutralizing murine and chimeric antibodies, fragments and regions of this invention.

Monoclonal antibodies of the present invention may be administered by any means that enables the active agent to reach the agent's site of action in the body of a mammal. Because proteins are subject to being digested when administered orally, parenteral administration, i.e., intravenous, subcutaneous, intramuscular, would ordinarily be used to optimize absorption.

Monoclonal antibodies of the present invention may be administered either as individual therapeutic agents or in combination with other therapeutic agents. They can be administered alone, but are generally administered with a pharmaceutical carrier selected on the basis of the chosen route of administration and standard pharmaceutical practice.

The dosage administered will, of course, vary depending upon known factors such as the pharmacodynamic characteristics of the particular agent, and its mode and route of administration; age, health, and weight of the recipient; nature and extent of symptoms, kind of concurrent treatment, frequency of treatment, and the effect desired. Usually a daily dosage of active ingredient can be about 0.1 to 100 milligrams per kilogram of body weight. Ordinarily 0.5 to 50, and preferably 1 to 10 milligrams per

kilogram per day given in divided doses 1 to 6 times a day or in sustained release form is effective to obtain desired results.

Dosage forms (composition) suitable for internal administration generally contain from about 1 milligram to about 500 milligrams of active ingredient per unit. In these pharmaceutical compositions the active ingredient will ordinarily be present in an amount of about 0.5-95% by weight based on the total weight of the composition.

For parenteral administration, the antibody can be formulated as a solution, suspension, emulsion or lyophilized powder in association with a pharmaceutically acceptable parenteral vehicle. Examples of such vehicles are water, saline, Ringer's solution, dextrose solution, and 5% human serum albumin. Liposomes and nonaqueous vehicles such as fixed oils may also be used. The vehicle or lyophilized powder may contain additives that maintain isotonicity (e.g., sodium chloride, mannitol) and chemical stability (e.g., buffers and preservatives). The formulation is sterilized by commonly used techniques.

Suitable pharmaceutical carriers are described in the most recent edition of Remington's Pharmaceutical Sciences, A. Osol, a standard reference text in this field.

The murine and chimeric antibodies, fragments and regions of this invention, their fragments, and derivatives can be used therapeutically as immunoconjugates (see for review: Dillman, R.O., *Ann. Int. Med.* 111:592-603 (1989)). They can be coupled to cytotoxic proteins, including, but not limited to Ricin-A, Pseudomonas toxin, and Diphtheria toxin. Toxins conjugated to antibodies or other ligands, are known in the art (see, for example, Olsnes, S. *et al.*, *Immunol. Today* 10:291-295 (1989)). Plant and bacterial toxins typically kill cells by disrupting the protein synthetic machinery.

The antibodies of this invention can be conjugated to additional types of therapeutic moieties including, but not limited to, radionuclides, cytotoxic agents and drugs. Examples of radionuclides which can be coupled to antibodies and delivered *in vivo* to sites of antigen include
5 ^{212}Bi , ^{131}I , ^{186}Re , and ^{90}Y , which list is not intended to be exhaustive. The radionuclides exert their cytotoxic effect by locally irradiating the cells, leading to various intracellular lesions, as is known in the art of radiotherapy.

Cytotoxic drugs which can be conjugated to antibodies and subsequently used for *in vivo* therapy include, but are not limited to,
10 daunorubicin, doxorubicin, methotrexate, and Mitomycin C. Cytotoxic drugs interfere with critical cellular processes including DNA, RNA, and protein synthesis. For a fuller exposition of these classes of drugs which are known in the art, and their mechanisms of action, see Goodman, A.G., *et al.*, *Goodman and Gilman's THE PHARMACOLOGICAL BASIS OF*
15 *THERAPEUTICS*, 7th Ed., Macmillan Publishing Co., 1985.

The antibodies of this invention may be advantageously utilized in combination with other monoclonal or murine and chimeric antibodies, fragments and regions, or with lymphokines or hemopoietic growth factors, etc., which serve to increase the number or activity of effector
20 cells which interact with the antibodies.

XII. Transgenic PCA3 Non-human Animals

Methods of Generating Transgenic Non-Human Animals

The non-human animals of the invention comprise any
25 animal having a transgenic interruption or alteration of the endogenous gene(s) (knock-out animals) and/or into the genome of which has been introduced one or more transgenes that direct the expression of human PCA3. Also preferred are the introduction of antisense PCA3 nucleic acids.

Such non-human animals include vertebrates such as rodents, non-human primates, sheep, dog, cow, amphibians, reptiles, etc. Preferred non-human animals are selected from non-human mammalian species of animals, most preferably, animals from the rodent family including rats and mice, most preferably mice.

The transgenic animals of the invention are animals into which has been introduced by nonnatural means (i.e., by human manipulation), one or more genes that do not occur naturally in the animal, e.g., foreign genes, genetically engineered endogenous genes, etc. The nonnaturally introduced genes, known as transgenes, may be from the same or a different species as the animal but not naturally found in the animal in the configuration and/or at the chromosomal locus conferred by the transgene. Transgenes may comprise foreign DNA sequences, i.e., sequences not normally found in the genome of the host animal. Alternatively or additionally, transgenes may comprise endogenous DNA sequences that are abnormal in that they have been rearranged or mutated *in vitro* in order to alter the normal *in vivo* pattern of expression of the gene, or to alter or eliminate the biological activity of an endogenous gene product encoded by the gene. (Watson, J.D., *et al.*, in *Recombinant DNA*, 2d Ed., W.H. Freeman & Co., New York (1992), pages 255-272; Gordon, J.W., *Intl. Rev. Cytol.* 115:171-229 (1989); Jaenisch, R., *Science* 240:1468-1474 (1989); Rossant, J., *Neuron* 2:323-334 (1990)).

The transgenic non-human animals of the invention are produced by introducing transgenes into the germline of the non-human animal. Embryonic target cells at various developmental stages are used to introduce the transgenes of the invention. Different methods are used depending on the stage of development of the embryonic target cell(s).

1. Microinjection of zygotes is the preferred method for incorporating transgenes into animal genomes in the course of practicing the

invention. A zygote, a fertilized ovum that has not undergone pronuclei fusion or subsequent cell division, is the preferred target cell for microinjection of transgenic DNA sequences. The murine male pronucleus reaches a size of approximately 20 micrometers in diameter, a feature which allows for the reproducible injection of 1-2 picoliters of a solution containing transgenic DNA sequences. The use of a zygote for introduction of transgenes has the advantage that, in most cases, the injected transgenic DNA sequences will be incorporated into the host animal's genome before the first cell division (Brinster, *et al.*, *Proc. Natl. Acad. Sci. (USA)* 82:4438-4442 (1985)). As a consequence, all cells of the resultant transgenic animals (founder animals) stably carry an incorporated transgene at a particular genetic locus, referred to as a transgenic allele. The transgenic allele demonstrates Mendelian inheritance: half of the offspring resulting from the cross of a transgenic animal with a non-transgenic animal will inherit the transgenic allele, in accordance with Mendel's rules of random assortment.

2. Viral integration can also be used to introduce the transgenes of the invention into an animal. The developing embryos are cultured *in vitro* to the developmental stage known as a blastocyst. At this time, the blastomeres may be infected with appropriate retroviruses (Jaenich, R., *Proc. Natl. Sci. (USA)* 73:1260-1264 (1976)). Infection of the blastomeres is enhanced by enzymatic removal of the zona pellucida (Hogan, *et al.*, in *Manipulating the Mouse Embryo*, Cold Spring Harbor Press, Cold Spring Harbor, N.Y. (1986)). Transgenes are introduced via viral vectors which are typically replication-defective but which remain competent for integration of viral-associated DNA sequences, including transgenic DNA sequences linked to such viral sequences, into the host animal's genome (Jahner, *et al.*, *Proc. Natl. Acad. Sci. (USA)* 82:6927-6931 (1985); Van der Putten, *et al.*, *Proc. Natl. Acad. Sci. (USA)* 82:6148-6152 (1985)). Transfection is easily and efficiently obtained by culture of blastomeres on a mono-layer of cells

producing the transgene-containing viral vector (Van der Putten, *et al.*, *Proc. Natl. Acad. Sci. (USA)* 82:6148-6152 (1985); Stewart, *et al.*, *EMBO Journal* 6:383-388 (1987)). Alternatively, infection may be performed at a later stage, such as a blastocoele (Jahner, D., *et al.*, *Nature* 298:623-628 (1982)). In any event, most transgenic founder animals produced by viral integration will be mosaics for the transgenic allele; that is, the transgene is incorporated into only a subset of all the cells that form the transgenic founder animal. Moreover, multiple viral integration events may occur in a single founder animal, generating multiple transgenic alleles which will segregate in future generations of offspring. Introduction of transgenes into germline cells by this method is possible but probably occurs at a low frequency (Jahner, D., *et al.*, *Nature* 298:623-628 (1982)). However, once a transgene has been introduced into germline cells by this method, offspring may be produced in which the transgenic allele is present in all of the animal's cells, i.e., in both somatic and germline cells.

3. Embryonic stem (ES) cells can also serve as target cells for introduction of the transgenes of the invention into animals. ES cells are obtained from pre-implantation embryos that are cultured *in vitro* (Evans, M.J., *et al.*, *Nature* 292:154-156 (1981); Bradley, M. O., *et al.*, *Nature* 309:255-258 (1984); Gossler, *et al.*, *Proc. Natl. Acad. Sci. (USA)* 83:9065-9069 (1986); Robertson *et al.*, *Nature* 322:445-448 (1986); Robertson, E.J., in *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach*, Robertson, E.J., ed., IRL Press, Oxford (1987), pages 71-112). ES cells, which are commercially available (from, e.g., Genome Systems, Inc., St. Louis, MO), can be transformed with one or more transgenes by established methods (Lovell-Badge, R.H., in *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach*, Robertson, E.J., ed., IRL Press, Oxford (1987), pages 153-182). Transformed ES cells can be combined with an animal blastocyst, whereafter the ES cells colonize the embryo and contribute to the

germline of the resulting animal, which is a chimera (composed of cells derived from two or more animals) (Jaenisch, R., *Science* 240:1468-1474 (1988); Bradley, A., in *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach*, Robertson, E.J., ed., IRL Press, Oxford (1987), pages 5 113-151). Again, once a transgene has been introduced into germline cells by this method, offspring may be produced in which the transgenic allele is present in all of the animal's cells, i.e., in both somatic and germline cells.

However it occurs, the initial introduction of a transgene is a Lamarckian (non-Mendelian) event. However, the transgenes of the 10 invention may be stably integrated into germ line cells and transmitted to offspring of the transgenic animal as Mendelian loci. Other transgenic techniques result in mosaic transgenic animals, in which some cells carry the transgenes and other cells do not. In mosaic transgenic animals in which germ line cells do not carry the transgenes, transmission of the transgenes 15 to offspring does not occur. Nevertheless, mosaic transgenic animals are capable of demonstrating phenotypes associated with the transgenes.

Transgenes may be introduced into non-human animals in order to provide animal models for human diseases. Transgenes that result in such animal models include, e.g., transgenes that encode mutant gene 20 products associated with an inborn error of metabolism in a human genetic disease and transgenes that encode a human factor required to confer susceptibility to a human pathogen (i.e., a bacterium, virus, or other pathogenic microorganism) (Leder *et al.*, U.S. Patent 5,175,383 (Dec. 29, 1992); Kindt *et al.*, U.S. Patent 5,183,949 (Feb. 2, 1993); Small *et al.*, *Cell* 25 46:13-18 (1986); Hooper *et al.*, *Nature* 326:292-295 (1987); Stacey *et al.*, *Nature* 332:131-136 (1988); Windle *et al.*, *Nature* 343:665-669 (1990); Katz *et al.*, *Cell* 74:1089-1100 (1993)). Transgenically introduced mutations comprise null ("knock-out") alleles in which a DNA sequence encoding a selectable and/or detectable marker is substituted for a genetic sequence

normally endogenous to a non-human animal. Resultant transgenic non-human animals that are predisposed to a disease, or in which the transgene causes a disease, may be used to identify compositions that induce the disease and to evaluate the pathogenic potential of compositions known or suspected to induce the disease (Berns, A.J.M., U.S. Patent 5,174,986 (Dec. 29, 1992)), or to evaluate compositions which may be used to treat the disease or ameliorate the symptoms thereof (Scott *et al.*, WO 94/12627 (1994)).

Offspring that have inherited the transgenes of the invention are distinguished from littermates that have not inherited transgenes by analysis of genetic material from the offspring for the presence of biomolecules that comprise unique sequences corresponding to sequences of, or encoded by, the transgenes of the invention. For example, biological fluids that contain polypeptides uniquely encoded by the selectable marker of the transgenes of the invention may be immunoassayed for the presence of the polypeptides. A more simple and reliable means of identifying transgenic offspring comprises obtaining a tissue sample from an extremity of an animal, e.g., a tail, and analyzing the sample for the presence of nucleic acid sequences corresponding to the DNA sequence of a unique portion or portions of the transgenes of the invention, such as the selectable marker thereof. The presence of such nucleic acid sequences may be determined by, e.g., hybridization ("Southern") analysis with DNA sequences corresponding to unique portions of the transgene, analysis of the products of PCR reactions using DNA sequences in a sample as substrates and oligonucleotides derived from the transgene's DNA sequence, etc.

The present invention is described in further detail in the following non-limiting examples.

EXAMPLE 1

Isolation and Characterization of PCA3 cDNA and Genomic DNA

In order to identify new markers for prostate cancer, a differential display analysis (Liang *et al.*, *Science* 257: 967-971 (1992)) was used to identify genes overexpressed in prostatic carcinomas in comparison to normal prostate; total RNA from normal, benign hyperplastic and malignant prostatic tissue from the same patients was extracted. Using twenty different combinations of primers (four anchored primers, five arbitrary primers), eleven apparently differentially expressed mRNAs were identified (i.e., consistently overexpressed in all carcinomas studied and not expressed in normal or BPH tissue). The complementary DNA (cDNA) fragments were used as probes for Northern blot analysis to confirm the consistent overexpression in the prostatic tumors used for the differential display. One of the probes (named DD3, a 486 bp cDNA) detected two major transcripts (2.3 and 4.0kb) that are highly overexpressed in 47 of 50 human prostatic tumors studied, whereas no (or very low levels of) expression of these transcripts was found in normal or BPH tissue from the same patients.

To obtain a full length cDNA clone, a cDNA library was constructed using mRNA isolated from human primary prostatic tumor tissue. 250 positive DD3 related clones were obtained from screening this library. 80 clones were purified and the nucleotide sequence of these clones was determined by automated sequence analysis.

A genomic library constructed of human placenta genomic DNA cloned in λ FIX2 was screened using DD3 as a probe. Four different clones were obtained, two of them located towards the 5' end of the gene (λ FIX-ME3 and -ME4) and two clones located towards the 3' end of the gene (λ FIX-ME1 and -ME2). The 5' end of λ FIX-ME4 was subcloned and used as a probe to screen the genomic library. Three new, unique clones were isolated (λ FIX-IH1, IH2, and IH6).

From the 80 analyzed cDNA clones, at least four different transcripts were shown to be present due to alternative splicing or alternative polyadenylation. Sequence analysis of the genomic clones as compared to the cDNA clones revealed the genomic structure of the PCA3 gene. Three
5 introns and 4 exons are present. The first intron is approximately 20 kb in length.

The first cDNA species is found in approximately 5% of the cDNA clones and contains exons 1, 2, 3, 4a and 4b (poly-adenylation after 4b is preceded by a real consensus poly-A-addition signal)(Figure 1).

10 The second cDNA species, found in approximately 15% of the cDNA cloned, contains exons 1, 3, 4a, 4b and 4c, arises by alternative splicing of the second exon (not present in this cDNA) and terminates at a different (real consensus) poly-A-addition signal (Figure 1).

The third cDNA species contains exons 1, 3, 4a, and 4b
15 and is the most common one found (approximately 65% of 80 clones) (Figure 1). This cDNA is most likely responsible for the most prominent transcript seen by Northern blot analysis (2 kb).

The fourth cDNA species detected contains exons 1, 3, and 4a representing about 15% of clones, and terminates after 4a, which is
20 the original DD3 clone stop site (Figure 1). The poly-A-addition signal present here is close to the consensus sequence.

PCA3 is a gene wherein significant alternative splicing (as well as alternative poly-adenylation) occurs, as evidenced by the differently sized transcripts observed on Northern blots and the different types of clones
25 identified. As mentioned previously, other splicing variants can be identified, as virtually every combination of exons is possible. For instance, a cDNA clone having exons 2, 3, 4a, 4b, and 4c has recently been identified. Indeed, it appears that clones representing virtually all possible exon combinations have been isolated.

One such splicing variant was identified by sequencing a clone named λ DD3.6. λ DD3.6 is a λ gt11 clone identified and isolated upon screening of a cDNA library made from prostate RNA of a 25 year old male (obtained from Clontech) with a PCA3 probe. λ DD3.6 contains exon 3, 4a, 4b, 4c, and 4d. However, this cDNA clone also contains intron sequences (part of intron 2, as well as intron 3).

A comparison of the two deposited clones PMB9 and λ DD3.6 is shown in Figure 3.

Different combinations of exons were examined by computer analysis to identify open reading frames (ORFs) and to predict the protein encoding region. The longest ORF was also the most highly likely protein encoding region. The longest ORF of 153 nucleotides encodes a small peptide of 51 amino acids, PCA3. PCA3 is encoded by part of exon 3 and 4a. The small size of the protein suggests that the protein most likely functions as a messenger molecule and has the potential to be secreted from the cells. The nucleotide sequence of exons 1-4a-d and the amino acid sequence of PCA3 are shown in Fig. 2 and Fig. 5 (SEQ ID NO:1 and 6 and 2 and 7, respectively).

It will be recognized by the person of ordinary skill, that a cDNA clone comprising the nucleic acid sequence presented in SEQ ID NO:6 and shown in Fig. 5 can be obtained as previously described by isolating and characterizing PCA3 cDNA clones. For example, and as commonly known in the art, probes which are specific to at least one of the 5' end, exon 1, 2, 3, 4a, 4b, 4c and 4d can be further used to increase the probability of having a full-length PCA3 cDNA clone. 96-well plates, for example, can be used to screen a large number of PCA3 positive cDNA clones, using the probes mentioned above. Of course, PCA3 positive clones can also be sequenced,

as commonly known and as described herein, until a desired cDNA clone is obtained.

In addition it is also possible to obtain a cDNA clone comprising the sequence shown in SEQ ID NO:6 and shown in Fig. 5, using PCA3 specific primers and an amplifying method such as PCR. For example, PCR technology with primers specific for the ultimate 5' and 3' end of the PCA3 cDNA, could be used to amplify a desired product (almost 4 kb) from RNA, isolated for example from prostatic tumors, and clone the PCR products. However, since PCR amplification may introduce mistakes, a sequencing of the complete cDNA would most likely be required.

As well known to the person of ordinary skill, a cDNA clone comprising the sequence shown in SEQ ID NO:6 and shown in Fig. 5 can also be constructed using the clones described herein (or newly isolated ones) and conventional genetic engineering methods.

For example, such a full length cDNA clone can be constructed using the deposited clones pMB9 and λ DD3.6. A non-limiting example of a strategy to construct such a cDNA clone comprising the nucleic acid sequence of SEQ ID NO:6 and Fig. 5, is described below.

λ DD3.6 phage DNA is digested to completion with *Nde*I, and the approximately 2 kb *Nde*I fragment isolated from an agarose gel. This fragment contains part of PCA3 exon 4b, exons 4c + 4d and about 50 nucleotides of phage DNA. The ends of this 2 kb fragment are then filled in with Klenow-fragment DNA polymerase and dNTPs, the blunt-ended fragment are then ligated into the *Hinc*II/*Sma*I sites of Bluescript SK. The loss of the *Hind*III site of Bluescript by the *Hinc*II and *Sma*I digestion is essential for further cloning steps in this particular strategy (see below). It is to be noted that *Nde*I sites are also present in phage λ gt11, giving rise to several additional fragments, some of which are close to 2 kb (*i.e.*, a 1.8 kb and a 2.5 kb fragment). Nevertheless, it is straight forward to separate these different bands on an agarose gel. The

correct orientation of insertion of the blunt-ended 2 kb *Nde*I fragment of λ DD3.6 into Bluescript (termed construct PCA3-X) can be verified by a single *Sac*I digestion, which should yield a ~0.45 and ~4.5 kb fragment by Ethidium-bromide staining of agarose gels. Sequence analysis may be performed to confirm the identity of the PCA3 insert.

The PCA3-X construct is then digested to completion with *Hind*III and *Bam*HI and a 4.8 kb vector-insert isolated from an agarose gel. This results in the removal of ~0.2 kb of DNA from the insert. pMB9 can be simultaneously digested to completion with *Bam*HI and *Hind*III and the 1.9 kb insert (containing PCA3 exons 1, 2, 3, 4a and most of exon 4b) isolated from an agarose gel. The pMB9-derived insert is ligated into the *Bam*HI/*Hind*III site of the PCA3-X construct. The resulting construct, PCA3-Y contains the complete cDNA of PCA3, except for the first 22 nucleotides of exon1 (see below and Fig. 4). These 22 nucleotides can be added to the PCA3 cDNA by cutting the PCA3-Y construct and the oligo-(74)-mer (SEQ ID NO:8) to completion with *Bam*HI and *Pst*I and ligating the oligomer in construct PCA3-Y, resulting in construct PCA3-Z. Nucleotide sequence analysis can be performed to verify that the oligo was properly ligated (i.e. to confirm that just one oligo was ligated and not a whole array of oligos). Of course, a sequencing of the resulting cDNA in PCA3-Z, can be performed to verify the integrity of the nucleic acid sequence.

Screening of a somatic cell hybrid panel revealed that the gene encoding PCA3 is located on human chromosome 9. Using a mixture of four PCA3-related genomic clones as a probe to hybridize to metaphase chromosomes of human lymphocytes, PCA3 was mapped to 9q21-22 (See also, Figure 1).

The conservation of PCA3 gene during evolution was studied by Southern blot analysis and revealed that a homolog of this gene is present in monkey, cow, horse, sheep, goat and pig. The gene is also

present in dog and cat. By comparison, the gene encoding PSA is only found in primates.

EXAMPLE 2

5 Prostate specific expression of PCA3

Upon developing PCA3 specific primers, RT-PCR analysis was performed using RNA from several normal human tissues. At 40 cycles of PCR, PCA3 related products in normal prostate and BPH tissues were amplified. PCA3 expression is very prostate specific since no PCA3 product
10 could be amplified under these conditions in the following normal human tissues: artery, brain, breast, bladder, colon, duodenum, heart, liver, lung, ovary, pancreas, placenta, seminal vesicles, skeletal muscle, skin, spinal cord, spleen and testis. Also in the human prostate cancer cell lines ALVA-31, DU145, JCA-1, PPC-1, PC3, and TSU-Pr1 no PCA3 related PCR product
15 could be detected. In the cell line LNCaP a product can be obtained after 40 cycles of PCR (whereas under the same conditions a product can be obtained in prostatic tumors within 20 cycles). The technology used to assess the prostate specific expression of PCA3 can be adapted in a diagnostic test for prostate cancer. In addition, it can be adapted to the identification of the
20 prostatic origin of a metastase.

Furthermore, a semi-quantitative RT-PCR analysis to compare the expression of PCA3 to that of PSA (prostate-specific antigen) and PSM (prostate-specific membrane antigen) and to establish if PCA3-RT-PCR analysis can be used to distinguish malignant from benign prostatic
25 specimens was performed. After quantification of the RT-reaction, 10 ng of cDNA was used for the PCR reaction and as a control, beta-2 microglobulin was also examined. PCA3 products found allowed a clear distinction between benign and malignant specimens in 23 of 25 cases studied whereas PSA and PSM could not make this distinction: approximately equal amounts of product

were found in normal and tumor samples. The expression of PSA and PCA3 was also compared by Northern blot analysis, which clearly shows the higher tumor-specificity of PCA3. At least a 20-fold overexpression of PCA3 in prostatic carcinomas as compared to normal or BPH tissues is observed.

5 This is distinctly different from expression of PSM and PSA, both of which are decreased in malignant versus benign tissues. Thus, PCA3 appears to be a good marker for diagnostic of prostate cancer.

An ideal tumor marker for prostate cancer should not only be able to positively distinguish between benign and malignant tissues but
10 also be able to predict clinical outcome (cure or progression) of patients afflicted with this disease. Data has shown that indeed, the level of expression of PCA3 tends to be positively correlated with tumor grade.

RISH (and eventually immunohistochemistry) is used to establish whether or not there is correlation between overexpression of PCA3,
15 tumor grade, stage, and clinical outcome. For both the paraffin-embedded and frozen specimens, long-term clinical follow-up is available. Using computer-assisted image analysis, quantitation of PCA3 expression levels as detected by RISH is performed and this is normalized to an external reference (Tamimi *et al.*, *Cancer Res.* 53: 5512-16 (1993); Tamimi *et al.*, *B.J.*
20 *Cancer* (1996)). A multivariate regression analysis including Gleason: grade, pathological tumor stage, clinical tumor stage, PSA levels and PCA3 expression will be used to establish whether PCA3 is an accurate predictor of progression and has (additional) prognostic value.

Reverse transcriptase polymerase chain reaction (RT-
25 PCR) assays have been developed to detect occult hematogenous micrometastatic cells that might otherwise have gone undetected by presently available staging modalities. Such RT-PCR assays have already been performed in patients with prostate cancer and other malignancies. A highly sensitive (nested) RT-PCR assay (or other types of amplification

assays including without being limited to NASBA, PCR, QB rep., SOA, TMA, and LCR (Winn-Deen, *J. Clin., Liquid Assay* 19: 21-26 (1996)) can be used to detect prostate cancer cells in the circulating blood of prostate cancer patients to identify patients at risk for having or developing metastases.

- 5 Experiments will include appropriate controls (e.g. β -2-microglobulin) and will be performed in a semi-quantitative way (i.e., quantify the cDNA synthesis and use equal amounts of input for the PCR analysis).

- 10 The molecular staging studies will be performed in the larger context of the BIOMED II program (*Markers for Prostate Cancer*). In this extensive collaborative study, PSA and PSM will be studied as well as other potentially interesting markers for prostate cancer. Blood samples are already being collected from patients that are diagnosed with prostatic disease in the participating institutions. An optimization of the collecting and handling of blood samples from patients for the detection of circulating tumor cells has been initiated. The use for example of vacutainer™ CPT-tubes (BecktonDickinson) for blood collection and purification of peripheral blood leukocytes in combination with a Trizol™ RNA-extraction procedure (guanidinium thiocyanate based) resulted in the preparation of RNA qualitatively and quantitatively suitable for PCR analysis. The use of PCA3-specific primers, to amplify PCA3 transcripts in RNA extracted from blood from prostate cancer patients, revealed that the presence of prostate cancer cells in the blood circulation, of not only patients with proven metastases, but also of patients with assumed localized disease, could be detected. More extensive studies on a larger patient population and a correlation with clinical data and follow-up will be carried out to determine the prognostic value of PCA3 for the individual prostate cancer patient.
- 25

Nested RT-PCR analysis (or similar amplification methods) should prove instrumental in determining whether there are any organs (not yet tested) that express PCA3. For example, Cowper's gland (same

embryonic origin as the prostate) and also skene's gland (female "homologue" to the prostate) will be tested for PCA3.

In one "normal" prostate tissue specimen that contained 10% of tumor cells, PCA3 expression was detected indicating the high sensitivity of PCA3 as a tumor marker. In this manner, PCA3 expression was also detected in a few BPH samples that were subsequently found to contain small areas of tumor cells. The level of expression of PCA3 in prostatic cancers shows a trend towards a positive correlation with tumor grade. These data are based on analysis of autoradiographs resulting from Northern blot hybridization.

The observation that PCA3 expression seems to increase with loss of differentiation is different from what is reported for PSA, since PSA expression levels decrease with loss of differentiation (Hakalahti *et al.*, *Int. J. Cancer* 55:590-597 (1993)). There is at least 20-fold overexpression of PCA3 in prostatic carcinomas in comparison to normal or BPH tissues. This is distinctly different from the expression of PSA which is reported to decrease in malignant versus benign tissues. PCA3 expression was detected in 4 of 4 metastases studied.

EXAMPLE 3

Identification of a transcription start site of PCA3

In order to determine the transcription start site of PCA3 primer extension analysis, S1-nuclease mapping and 5'RACE (rapid amplification of cDNA ends) assays were performed. The major transcription start site was found to be located within a range of 4 nucleotides (Figure 4).

The results of these experiments further lengthen the size of the cDNA in a 5' direction by a further 22 nt with respect to the cDNA sequence of pMB9 (SEQ ID NO:1 and Figure 2). This additional 5' polynucleotide sequence is also shown in SEQ ID NO:6 and Figure 5)

* * * * *

All publications mentioned hereinabove are hereby incorporated in their entirety by reference.

- 5 While the foregoing invention has been described in some detail for purposes of clarity and understanding, it will be appreciated by one skilled in the art from a reading of this disclosure that various changes in form and detail can be made without departing from the true scope of the invention and appended claims.

09402713.051300
00ET90"ET220460

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

- (A) NAME: BUSSEMAKERS, Marion J.G.
(B) STREET: St Geertruidestraat 27
(C) CITY: 6521 KG Nijmegen
(E) COUNTRY: The Netherlands
(F) POSTAL CODE (ZIP): 6521 KG

(ii) TITLE OF INVENTION: PCA3, PCA3 GENES, AND METHODS OF USE

(iii) NUMBER OF SEQUENCES: 8

(iv) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(vi) PRIOR APPLICATION DATA:

- (A) APPLICATION NUMBER: US 60/041,836
(B) FILING DATE: 10-APR-1997

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2037 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
(B) LOCATION: 379..531

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

AGAAGCTGGC ATCAGAAAAA CAGAGGGGAG ATTTGTGTGG CTGCAGCCGA GGGAGACCAG 60
GAAGATCTGC ATGGTGGGAA GGACCTGATG ATACAGAGGA ATTACAACAC ATATACTTAG 120
TGTTTCAATG AACACCAAGA TAAATAAGTG AAGAGCTAGT CCGCTGTGAG TCTCCTCAGT 180

Sub A1

0040241-061300

0040241-061300

GACACAGGGC TGGATCACCA TCGACGGCAC TTTCTGAGTA CTCAGTGCAG CAAAGAAAGA 240

CTACAGACAT CTCAATGGCA GGGGTGAGAA ATAAGAAAGG CTGCTGACTT TACCATCTGA 300

GGCCACACAT CTGCTGAAAT GGAGATAATT AACATCACTA GAAACAGCAA GATGACAATA 360

TAATGTCTAA GTAGTGAC ATG TTT TTG CAC ATT TCC AGC CCC TTT AAA TAT 411
Met Phe Leu His Ile Ser Ser Pro Phe Lys Tyr
1 5 10

CCA CAC ACA CAG GAA GCA CAA AAG GAA GCA CAG AGA TCC CTG GGA GAA 459
Pro His Thr Gln Glu Ala Gln Lys Glu Ala Gln Arg Ser Leu Gly Glu
15 20 25

ATG CCC GGC CGC CAT CTT GGG TCA TCG ATG AGC CTC GCC CTG TGC CTG 507
Met Pro Gly Arg His Leu Gly Ser Ser Met Ser Leu Ala Leu Cys Leu
30 35 40

GTC CCG CTT GTG AGG GAA GGA CAT TAGAAAATGA ATTGATGTGT TCCTTAAAGG 561
Val Pro Leu Val Arg Glu Gly His
45 50

ATGGGCAGGA AAACAGATCC TGTGTGGAT ATTTATTTGA ACGGGATTAC AGATTTGAAA 621

TGAAGTCACA AAGTGAGCAT TACCAATGAG AGGAAAACAG ACGAGAAAAT CTTGATGGCT 681

TCACAAGACA TGCAACAAAC AAAATGGAAT ACTGTGATGA CATGAGGCAG CCAAGCTGGG 741

GAGGAGATAA CCACGGGGCA GAGGGTCAGG ATTCTGGCCC TGCTGCCTAA ACTGTGCGTT 801

CATAACCAAA TCATTTTATA TTTCTAACCC TCAAAACAAA GCTGTTGTAA TATCTGATCT 861

CTACGGTTCC TTCTGGGCCC AACATTCTCC ATATATCCAG CCACACTCAT TTTTAATATT 921

TAGTTCCCAG ATCTGTACTG TGACCTTTCT ACACTGTAGA ATAACATTAC TCATTTTGTT 981

CAAAGACCCT TCGTGTGCT GCCTAATATG TAGCTGACTG TTTTTCCTAA GGAGTGTCTT 1041

GGCCCAGGGG ATCTGTGAAC AGGCTGGGAA GCATCTCAAG ATCTTTCCAG GGTTATACTT 1101

ACTAGCACAC AGCATGATCA TTACGGAGTG AATTATCTAA TCAACATCAT CCTCAGTGTC 1161

TTTGCCCATTA CTGAAATTCA TTTCCCACTT TTGTGCCCAT TCTCAAGACC TCAAAATGTC 1221

ATTCCATTAA TATCACAGGA TTAACTTTTT TTTTAAACCT GGAAGAATTC AATGTTACAT 1281

GCAGCTATGG GAATTTAATT ACATATTTTG TTTTCCAGTG CAAAGATGAC TAAGTCCTTT 1341

ATCCCTCCCC TTTGTTTGAT TTTTTCCTCA GTATAAAGTT AAAATGCTTA GCCTTGTAAT 1401

[illegible]

(2) INFORMATION FOR SEO \ ID NO: 2:

(A) LENGTH: 51 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Glu Gly His
50

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1872 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

AGAAGCTGGC ATCAGAAAAA CAGAGGGGAG ATTTGTGTGG CTGCAGCCGA GGGAGACCAG 60
GAAGATCTGC ATGGTGGGAA GGACCTGATG ATACAGAGGT GAGAAATAAG AAAGGCTGCT 120
GACTTTACCA TCTGAGGCCA CACATCTGCT GAAATGGAGA TAATTAACAT CACTAGAAAC 180
AGCAAGATGA CAATATAATG TCTAAGTAGT GACATGTTTT TGCACATTTC CAGCCCCTTT 240
AAATATCCAC ACACACAGGA AGCACAAAAG GAAGCACAGA GATCCCTGGG AGAAATGCCC 300
GGCCGCCATC TTGGGTCATC GATGAGCCTC GCCCTGTGCC TGGTCCCGCT TGTGAGGGAA 360
GGACATTAGA AAATGAATTG ATGTGTTTCT TAAAGGATGG GCAGGAAAAC AGATCCTGTT 420
GTGGATATTT ATTTGAACGG GATTACAGAT TTGAAATGAA GTCACAAAGT GAGCATTACC 480
AATGAGAGGA AAACAGACGA GAAAATCTTG ATGGCTTCAC AAGACATGCA ACAAACAAAA 540
TGGAATACTG TGATGACATG AGGCAGCCAA GCTGGGGAGG AGATAACCAC GGGGCAGAGG 600
GTCAGGATTC TGGCCCTGCT GCCTAAACTG TGCCTTCATA ACCAAATCAT TTCATATTTT 660
TAACCCTCAA AACAAAGCTG TTGTAATATC TGATCTCTAC GGTTCCTTCT GGGCCCAACA 720
TTCTCCATAT ATCCAGCCAC ACTCATTTTT AATATTTAGT TCCCAGATCT GTACTGTGAC 780
CTTTCTACAC TGTAGAATAA CATTACTCAT TTTGTTCAAA GACCCTTCGT GTTGCTGCCT 840
AATATGTAGC TGACTGTTTT TCCTAAGGAG TGTTCCTGGC CAGGGGATCT GTGAACAGGC 900
TGGAAGCAT CTCAAGATCT TTCCAGGGTT ATACTTAQTA GCACACAGCA TGATCATTAC 960
GGAGTGAATT ATCTAATCAA CATCATCCTC AGTGTCTTTC CCCATACTGA AATTCATTTC 1020
CCACTTTTGT GCCCATTCCTC AAGACCTCAA AATGTCATTTC CATTAAATATC ACAGGATTAA 1080
CTTTTTTTTT TAACCTGGAA GAATTCAATG TTACATGCAG CTATGGGAAT TTAATTACAT 1140

ATTTTGT	TTT	CCAGTGCAAA	GATGACTAAG	TCCTTTATCC	CTCCCCTTTG	TTTGATTTTT	1200
TTTCCAGTAT	AAAGTTAAAA	TGCTTAGCCT	TGTACTGAGG	CTGTATACAG	CACAGCCTCT		1260
CCCCATCCCT	CCAGCCTTAT	CTGTCATCAC	CATCAACCCC	TCCCATNYS	AA	CCTAAACAAA	1320
ATCTAACTTG	TAATTCCTTG	AACATGTCAG	GNCATACATT	RTTCCTTCTG	CCTGAGAAGC		1380
TCTTCCTTGT	CTCTTAANTC	TAGAATGATG	TAAAGTTTTG	AATAAGTTGA	CTATCTTACT		1440
TCATGCAAAG	AAGGGACACA	TATGAGATT	CATCATCACAT	GAGACAGCAA	ATACTAAAAG		1500
TGTAATTTGA	TTATAAGAGT	TTAGATAAAT	ATATGAAATG	CAAGAKCCAC	AGAGGGAATG		1560
TTTATGGGGC	ACGTTTGTAA	GCCTGGGATG	TGAAGMAAAG	GCAGGGAACC	TCATAGTATC		1620
TTATATAATA	TACTTCATTT	CTCTATCTCT	ATCACAATAT	CCAACAAGCT	TTTCACAGAA		1680
TTCATGCAGT	GCAAATCCCC	AAAGGTAACC	TTTATCCATT	TCATGGTGAG	TGCGCTTTAG		1740
AATTTTGGCA	AATCATACTG	GTCACTTATC	TCAACTTTGA	GATGTGTTTG	TCCTTGTAGT		1800
TAATTGAAAG	AAATAGGGCA	CTCTTGTGAG	CCACTTTAGG	GTTCACTCCT	GGCAATAAAG		1860
AATTTACAAA	GA						1872

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 820 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

AGAAGCTGGC	ATCAGAAAAA	CAGAGGGGAG	ATTTGTGTGG	CTGCAGCCGA	GGGAGACCAG	60
GAAGATCTGC	ATGGTGGGAA	GGACCTGATG	ATACAGAGGT	GAGAAATAAG	AAAGGCTGCT	120
GACTTTACCA	TCTGAGGCCA	CACATCTGCT	GAAATGGAGA	TAATTAACAT	CACTAGAAAC	180
AGCAAGATGA	CAATATAATG	TCTAAGTAGT	GACATGTTTT	TGCACATTTC	CAGCCCCTTT	240
AAATATCCAC	ACACACAGGA	AGCACAAAAG	GAAGCAGAGA	GATCCCTGGG	AGAAATGCCC	300
GGCCGCCATC	TTGGGTCATC	GATGAGCCTC	GCCCTGTGCC	TGGTCCCGCT	TGTGAGGGAA	360

Sub A1

004037.051300

```

GGACATTAGA AAATGAATTG ATGTGTTCTT TAAAGGATGG GCAGGAAAAC AGATCCTGTT      420
GTGGATATTT ATTTGAACGG GATTACAGAT TTGAAATGAA GTCACAAAGT GAGCATTACC      480
AATGAGAGGA AAACAGACGA GAAAATCTTG ATGGCTTCAC AAGACATGCA ACAAACAAAA      540
TGGAATACTG TGATGACATG AGGCAGCCAA GCTGGGGAGG AGATAACCAC GGGGCAGAGG      600
GTCAGGATTC TGGCCCTGCT GCCTAAACTG TGC GTTCATA ACCAAATCAT TTCATATTTT      660
TAACCCTCAA AACAAAGCTG TTGTAATATC TGATCTCTAC GGTTCCTTCT GGGCCCAACA      720
TTCTCCATAT ATCCAGCCAC ACTCATTTTT AATATTTAGT TCCCAGATCT GTACTGTGAC      780
CTTTCTACAC TGTAGAATAA CATTACTCAT TTTGTTCAAA      820

```

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 11 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

```

His Thr Gln Glu Ala Gln Lys Glu Ala Gln Arg
1           5           10

```

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3582 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 401..553

(xi) \ SEQUENCE DESCRIPTION: SEQ ID NO: 6:

ACAGAAGAAA	TAGCAAGTGC	CGAGAAGCTG	GCATCAGAAA	AACAGAGGGG	AGATTTGTGT		60
GGCTGCAGCC	GAGGGAGACC	AGGAAGATCT	GCATGGTGGG	AAGGACCTGA	TGATACAGAG		120
GAATTACAAC	ACATATACTT	AGTGTTCCTA	TGAACACCAA	GATAAATAAG	TGAAGAGCTA		180
GTCCGCTGTG	AGTCTCCTCA	GTGACACAGG	GCTGGATCAC	CATCGACGGC	ACTTTCTGAG		240
TACTCAGTGC	AGCAAAGAAA	GACTACAGAC	ATCTCAATGG	CAGGGGTGAG	AAATAAGAAA		300
GGCTGCTGAC	TTTACCATCT	GAGGCCACAC	ATCTGCTGAA	ATGGAGATAA	TTAACATCAC		360
TAGAAACAGC	AAGATGACAA	TATAATGTCT	AAGTAGTGAC	ATG TTT TTG CAC ATT			415
				Met Phe Leu His Ile			
					55		
TCC AGC CCC TTT AAA TAT CCA CAC ACA CAG GAA GCA CAA AAG GAA GCA							463
Ser Ser Pro Phe Lys Tyr Pro His Thr Gln Glu Ala Gln Lys Glu Ala							
	60			65		70	
CAG AGA TCC CTG GGA GAA ATG CCC GGC CGC CAT CTT GGG TCA TCG ATG							511
Gln Arg Ser Leu Gly Glu Met Pro Gly Arg His Leu Gly Ser Ser Met							
	75			80		85	
AGC CTC GCC CTG TGC CTG GTC CCG CTT GTG AGG GAA GGA CAT							553
Ser Leu Ala Leu Cys Leu Val Pro Leu Val Arg Glu Gly His							
	90			95		100	
TAGAAAATGA ATTGATGTGT TCCTTAAAGG ATGGGCAGGA AACAGATCC TGTTGTGGAT							613
ATTTATTTGA ACGGGATTAC AGATTTGAAA TGAAGTCACA AAGTGAGCAT TACCAATGAG							673
AGGAAAACAG ACGAGAAAAT CTTGATGGCT TCACAAGACA TGCAACAAAC AAAATGGAAT							733
ACTGTGATGA CATGAGGCAG CCAAGCTGGG GAGGAGATAA CCACGGGGCA GAGGGTCAGG							793
ATTCTGGCCC TGCTGCCTAA ACTGTGCGTT CATAACCAA TCATTTTCATA TTTCTAACCC							853
TCAAAACAAA GCTGTTGTAA TATCTGATCT CTACGGTTCC TTCTGGGCCC AACATTCTCC							913
ATATATCCAG CCACACTCAT TTTTAATATT TAGTTCCCAG ATCTGTACTG TGACCTTTCT							973
AACTGTAGA ATAACATTAC TCATTTTGTT CAAAGACCCT TCGTGTGCT GCCTAATATG							1033
TAGCTGACTG TTTTTCCTAA GGAGTGTTCCT GGCCCAGGGG ATCTGTGAAC AGGCTGGGAA							1093
GCATCTCAAG ATCTTTCCAG GGTATACTT ACTAGCACAC AGCATGATCA TTACGGAGTG							1153

Sub A1

00402713.06100

00402713.06100

AATTATCTAA	TCAACATCAT	CCTCAGTGTC	TTGCCCATA	CTGAAATTCA	TTTCCCACTT	1213
TTGTGCCCCAT	TCTCAAGACC	TCAAAATGTC	ATTCCATTAA	TATCACAGGA	TTAACTTTTT	1273
TTTTTAACCT	GGAAGAATTC	AATGTTACAT	GCAGCTATGG	GAATTTAATT	ACATATTTTG	1333
TTTTCCAGTG	CAAAGATGAC	TAAGTCCTTT	ATCCCTCCCC	TTTGTGTTGAT	TTTTTTTCCA	1393
GTATAAAGTT	AAAATGCTTA	GCCTTGTA	GAGGCTGTAT	ACAGCACAGC	CTCTCCCCAT	1453
CCCTCCAGCC	TTATCTGTCA	TCACCATCAA	CCCCTCCCAT	ACCACCTAAA	CAAAATCTAA	1513
CTTGTAATTC	CTTGAACATG	TCAGGACATA	CATTATTCCT	TCTGCCTGAG	AAGCTCTTCC	1573
TTGTCTCTTA	AATCTAGAAT	GATGTAAAGT	TTTGAATAAG	TTGACTATCT	TACTTCATGC	1633
AAAGAAGGGA	CACATATGAG	ATTCATCATC	ACATGAGACA	GCAAATACTA	AAAGTGTAAT	1693
TTGATTATAA	GAGTTTAGAT	AAATATATGA	AATGCAAGAG	CCACAGAGGG	AATGTTTATG	1753
GGGCACGTTT	GTAAGCCTGG	GATGTGAAGC	AAAGGCAGGG	AACCTCATAG	TATCTTATAT	1813
AATATACTTC	ATTTCTCTAT	CTCTATCACA	ATATCCAACA	AGCTTTTCAC	AGAATTCATG	1873
CAGTGCAAAT	CCCCAAAGGT	AACCTTTATC	CATTTTCATGG	TGAGTGCGCT	TTAGAATTTT	1933
GGCAAATCAT	ACTGGTCACT	TATCTCAACT	TTGAGATGTG	TTTGTCTTGG	TAGTTAATTG	1993
AAAGAAATAG	GGCACTCTTG	TGAGCCACTT	TAGGGTTCAC	TCCTGGCAAT	AAAGAATTTA	2053
CAAAGAGCTA	CTCAGGACCA	GTTGTTAAGA	GCTCTGTGTG	TGTGTGTGTG	TGTGTGTGAG	2113
TGTACATGCC	AAAGTGTGCC	TCTCTCTCTT	GACCATATAT	TTCAGACTTA	AAACAAGCAT	2173
GTTTTCAAAT	GGCACTATGA	GCTGCCAATG	ATGTATCACC	ACCATATCTC	ATTATCTCTC	2233
AGTAAATGTG	ATAATAATGT	CATCTGTAA	CATAAAAAAA	GTTTGACTTC	ACAAAAGCAG	2293
CTGGAAATGG	ACAACCACAA	TATGCATAAA	TCTAACTCCT	ACCATCAGCT	ACACACTGCT	2353
TGACATATAT	TGTTAGAAGC	ACCTCGCATT	TGTGGGTTCT	CTTAAGCAAA	ATACTTGCA	2413
TAGGTCTCAG	CTGGGGCTGT	GCATCAGGCG	GTTTGAGAAA	TATTCAATTC	TCAGCAGAAG	2473
CCAGAATTTG	AATCCCTCA	TCTTTTAGGA	ATCATTTACC	AGGTTTGAG	AGGATTCAGA	2533
CAGCTCAGGT	GCTTTCACTA	ATGTCTCTGA	ACTTCTGTCC	CTCTTTGTGT	TCATGGATAG	2593
TCCAATAAAT	AATGTTATCT	TTGAACTGAT	GCTCATAGGA	GAGAATATAA	GAACCTCTGAG	2653

Sub A1

09402713-051300

TGATATCAAC	ATTAGGGATT	CAAAGAAATA	TTAGATTTAA	GCTCACACTG	GTCAAAAGGA	2713
ACCAAGATAC	AAAGAACTCT	GAGCTGTCAT	CGTCCCCATC	TCTGTGAGCC	ACAACCAACA	2773
GCAGGACCCA	ACGCATGTCT	GAGATCCTTA	AATCAAGGAA	ACCAGTGTCA	TGAGTTGAAT	2833
TCTCCTATTA	TGGATGCTAG	CTTCTGGCCA	TCTCTGGCTC	TCCTCTTGAC	ACATATTAGC	2893
TTCTAGCCTT	TGCTTCCACG	ACTTTTATCT	TTTCTCCAAC	ACATCGCTTA	CCAATCCTCT	2953
CTCTGCTCTG	TTGCTTGGGA	CTTCCCCACA	AGAATTTCAA	CGACTCTCAA	GTCTTTTCTT	3013
CCATCCCCAC	CACTAACCTG	AATTGCCTAG	ACCCTTATTT	TTATTAATTT	CCAATAGATG	3073
CTGCCTATGG	GCTAATATTC	CTTTAGATGA	ACATTAGATA	TTTAAAGTCT	AAGAGGTTCA	3133
AAATCCAAC	CATTATCTTC	TCTTTCTTTC	ACCTCCCCTG	CTCCTCTCCC	TATATTACTG	3193
ATTGACTGAA	CAGGATGGTC	CCCAAGATGC	CAGTCAAATG	AGAAACCCAG	TGGCTCCTTG	3253
TGGATCATGC	ATGCAAGACT	GCTGAAGCCA	GAGGATGACT	GATTACGCCT	CATGGGTGGA	3313
GGGGACCACT	CCTGGGCCTT	CGTGATTGTC	AGGAGCAAGA	CCTGAGATGC	TCCCTGCCTT	3373
CAGTGTCCTC	TGCATCTCCC	CTTTCTAATG	AAGATCCATA	GAATTTGCTA	CATTTGAGAA	3433
TTCCAATTAG	GAACACACAT	GTTTTATCTG	CCCTATCAAT	TTTTTAAACT	TGCTGAAAAT	3493
TAAGTTTTTT	CAAAATCTGT	CCTTGTAAT	TACTTTTTTCT	TACAGTGTCT	TGGCATACTA	3553
TATCAACTTT	GATTCTTTGT	TACAACTTT				3582

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 51 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

Met	Phe	Leu	His	Ile	Ser	Ser	Pro	Phe	Lys	Tyr	Pro	His	Thr	Gln	Glu
1				5					10					15	
Ala	Gln	Lys	Glu	Ala	Gln	Arg	Ser	Leu	Gly	Glu	Met	Pro	Gly	Arg	His
			20						25					30	

Leu Gly Ser Ser Met Ser Leu Ala Leu Cys Leu Val Pro Leu Val Arg
35 40 45

Glu Gly His
50

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 74 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

GGATCCACAG AAGAAATAGC AAGTGCCGAG AAGCTGGCAT CAGAAAAACA GAGGGGAGAT 60

TTGTGTGGCT GCAG 74

81
~~83~~

Although the present invention has been described hereinabove by way of preferred embodiments thereof, it can be modified, without departing from the spirit and nature of the subject invention as defined in the appended claims.

09402713-061300